

S Claim 1; Page 103; 137pp; English.

C AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
C AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
C SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
C derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
C antibiotic, antimicrobial and antiviral activities, and can be used as
C microbial growth and proliferation inhibitors and in gene therapy. (1)
C are useful for inhibiting microbial growth in an environment capable of
C sustaining such growth, for inhibiting microbial growth or strain in a
C host, and inhibiting the growth of drug-resistant microbial strains such
C as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
C *Xanthomonas*.

Q Sequence 14 AA:

Query Match 100.0%; Score 66; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 LRRIRKIIHIKK 14
|||||
b 1 Lrrlrkllhikk 14

RESULT 2
AAB70659 standard; peptide; 14 AA.

X AAB70659;

T 15-MAY-2001 (first entry)

E Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:12.

M Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
M bactericidal; antibiotic; antiviral; microbial growth inhibitor;
M proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
M *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

X Ovis aries.

N WO200112668-A1.

X 22-FEB-2001.

D 18-AUG-2000; 2000WO-US22781.

F 18-AUG-1999; 99US-0149886.

R (IOWA) UNIV IOWA RES FOUND.
A (REGC) UNIV CALIFORNIA.

X Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
I WPI; 2001-234911/24.

X New antimicrobial peptides useful as antibiotics for inhibiting growth
T and proliferation of microbes, and for treating microbial infections -
X

S Claim 1; Page 103; 137pp; English.

C AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
C AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
C SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
C derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
C antibiotic, antimicrobial and antiviral activities, and can be used as
C microbial growth and proliferation inhibitors and in gene therapy. (1)
C are useful for inhibiting microbial growth in an environment capable of
C sustaining such growth, for inhibiting microbial growth or strain in a
C host, and inhibiting the growth of drug-resistant microbial strains such
C as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
C *Xanthomonas*.

CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.

XX

SO Sequence 14 AA:

Query Match 100.0%; Score 66; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRRIRKIIHIKK 14
|||||
Db 1 Lrrlrkllhikk 14

RESULT 3
AAB70656 standard; peptide; 16 AA.

ID AAB70656;

AC AAB70656;

XX 15-MAY-2001 (first entry)

DT

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:9.

XX

KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

XX

OS Ovis aries.

OS

PN WO200112668-A1.

XX

PD 22-FEB-2001.

XX

PF 18-AUG-2000; 2000WO-US22781.

PR 18-AUG-1999; 99US-0149886.

XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.

PA

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.

DR

XX

XX

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX

PS Claim 1; Page 103; 137pp; English.

XX

CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.

XX

SO Sequence 16 AA:

Query Match 100.0%; Score 66; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00035;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRRIRKIIHIKK 14
|||||

Db 2 lrrlrklhikk 15

RESULT 4
AAB70648
ID AAB70648 standard; peptide: 18 AA.
XX

AC AAB70648;
XX
DT 15-MAY-2001 (first entry)
XX

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:1.
XX

OS Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX

OS Ovis aries.
XX
PN MO200112668-A1.
XX
PD 22-FEB-2001.
XX

PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX

PA (IOWA) UNIV IOWA RES FOUND.
XX (REGC) UNIV CALIFORNIA.
XX

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections -
XX

PS Claim 1; Page 103; 137pp; English.
XX

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
XX AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
XX SMAP 29 cathelicidin family peptide, and AAB70673 are
XX derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
XX antibiotic, antimicrobial and antiviral activities, and can be used as
XX microbial growth and proliferation inhibitors and in gene therapy. (I)
XX are useful for inhibiting microbial growth in an environment capable of
XX sustaining such growth, for inhibiting microbial growth or strain in a
XX host, and inhibiting the growth of drug-resistant microbial strains such
XX as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX Xanthomonas.
XX

SQ Sequence 18 AA:
XX

QY 1 LRRIRKIIHIKK 14
XX
DB 3 lrrlrklhikk 16
XX

Query Match 100.0%; Score 66; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.00039;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
AAB70655
ID AAB70655 standard; peptide: 18 AA.
XX
AC AAB70655;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:8.
XX

XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX

OS Ovis aries.
XX
PN MO200112668-A1.
XX
PD 22-FEB-2001.
XX

PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX

PA (IOWA) UNIV IOWA RES FOUND.
XX (REGC) UNIV CALIFORNIA.
XX

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections -
XX

PS Claim 1; Page 103; 137pp; English.
XX

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
XX AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
XX SMAP 29 cathelicidin family peptide, and AAB70673 are
XX derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
XX antibiotic, antimicrobial and antiviral activities, and can be used as
XX microbial growth and proliferation inhibitors and in gene therapy. (I)
XX are useful for inhibiting microbial growth in an environment capable of
XX sustaining such growth, for inhibiting microbial growth or strain in a
XX host, and inhibiting the growth of drug-resistant microbial strains such
XX as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX Xanthomonas.
XX

SQ Sequence 18 AA:
XX

QY 1 LRRIRKIIHIKK 14
XX
DB 3 lrrlrklhikk 16
XX

Query Match 100.0%; Score 66; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.00039;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
AAB70675
ID AAB70675 standard; peptide: 29 AA.
XX
AC AAB70675;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:28.
XX

OS Ovis aries.
XX
PN MO200112668-A1.
XX
PD 22-FEB-2001.
XX

CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.
 XX
 SQ Sequence 14 AA;

Query Match 97.0%; Score 64; DB 22; Length 14;
 Best Local Similarity 92.9%; Pred. No. 0.00064;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LRRIRKIIHIKK 14
 :|||||
 Db 1 IRRIRKIIHIKK 14

RESULT 9
 AAB70657
 ID AAB70657 standard; peptide; 16 AA.
 XX
 AC AAB70657;
 XX

DT 15-MAY-2001 (first entry)
 XX

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:10.
 XX

KM Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
 KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
 KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
 XX

OS Ovis aries.
 XX

PN WO200112668-A1.
 XX

PD 22-FEB-2001.
 XX

PF 18-AUG-2000; 2000WO-US22781.
 XX

PR 18-AUG-1999; 99US-0149886.
 XX

PA (IOWA) UNIV IOWA RES FOUND.
 PA (REGC) UNIV CALIFORNIA.
 XX

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
 XX

DR WPI: 2001-234911/24.
 XX

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.
 CC
 XX

SQ Sequence 16 AA;

Query Match 97.0%; Score 64; DB 22; Length 16;
 Best Local Similarity 92.9%; Pred. No. 0.00072;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LRRIRKIIHIKK 14

Db :|||||
 2 IRRIRKIIHIKK 15

RESULT 10
 AAB70654
 ID AAB70654 standard; peptide; 18 AA.
 XX
 AC AAB70654;
 XX

DT 15-MAY-2001 (first entry)
 XX

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:7.
 XX

KM Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
 KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
 KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
 XX

OS Ovis aries.
 XX

PN WO200112668-A1.
 XX

PD 22-FEB-2001.
 XX

PF 18-AUG-2000; 2000WO-US22781.
 XX

PR 18-AUG-1999; 99US-0149886.
 XX

PA (IOWA) UNIV IOWA RES FOUND.
 PA (REGC) UNIV CALIFORNIA.
 XX

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
 XX

DR WPI: 2001-234911/24.
 XX

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.
 CC
 XX

SQ Sequence 18 AA;

Query Match 97.0%; Score 64; DB 22; Length 18;
 Best Local Similarity 92.9%; Pred. No. 0.00081;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LRRIRKIIHIKK 14
 :|||||
 Db 3 IRRIRKIIHIKK 16

RESULT 11

AAB70653
 ID AAB70653 standard; peptide; 20 AA.
 XX

AC AAB70653;
 XX

DT 15-MAY-2001 (first entry)
 XX

```

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:6.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
OS
XX Ovis aries.
XX WO200112668-A1.
XX
XX PD 22-FEB-2001.
XX
XX PF 18-AUG-2000; 2000WO-US22781.
XX
XX PR 18-AUG-1999; 99US-0149886.
XX
XX PA (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R:
XX WPI; 2001-234911/24.
XX
XX PS New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (1'), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
XX SQ Sequence 20 AA:
XX
XX
XX Query Match 97.0%; Score 64; DB 22; Length 20;
XX Best Local Similarity 92.9%; Pred. No. 0.00089;
XX Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX OY 1 LRRIRKIIHIKK 14
XX :|||||
XX 3 IRIIRKIIHIKK 16
XX
XX DB
XX
XX RESULT 12
XX AAB70663
XX ID AAB70663 standard; peptide; 13 AA.
XX
XX AC AAB70663;
XX
XX DT 15-MAY-2001 (first entry)
XX
XX DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:16.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX OS Ovis aries.
XX
XX XX WO200112668-A1.
XX
XX PD 22-FEB-2001.
XX

```

```

XX
XX 18-AUG-2000; 2000WO-US22781.
XX
XX PR 18-AUG-1999; 99US-0149886.
XX
XX PA (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R:
XX WPI; 2001-234911/24.
XX
XX PS New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (1'), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
XX SQ Sequence 13 AA:
XX
XX
XX Query Match 93.9%; Score 62; DB 22; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 0.0012;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX OY 2 RRIIRKIIHIKK 14
XX :|||||
XX 1 RRIIRKIIHIKK 13
XX
XX DB
XX
XX RESULT 13
XX AAB70662
XX ID AAB70662 standard; peptide; 13 AA.
XX
XX AC AAB70662;
XX
XX DT 15-MAY-2001 (first entry)
XX
XX DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:15.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX OS Ovis aries.
XX
XX XX WO200112668-A1.
XX
XX PD 22-FEB-2001.
XX
XX PF 18-AUG-2000; 2000WO-US22781.
XX
XX PR 18-AUG-1999; 99US-0149886.
XX
XX PA (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R:
XX WPI; 2001-234911/24.
XX
XX

```

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.

XX
SQ Sequence 13 AA;

Query Match 92.4%; Score 61; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LRIRKIIHIK 13
Db 1 LRIRKIIHIK 13

RESULT 14

AAB70664
ID AAB70664 standard; peptide; 12 AA.

XX
AC AAB70664;

XX
DT 15-MAY-2001 (first entry)

XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:17.

XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;

KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;

KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

XX
OS Ovis aries.

XX
PN WO200112668-A1.

XX
PD 22-FEB-2001.

XX
PF 18-AUG-2000; 2000WO-US22781.

XX
PR 18-AUG-1999; 99US-0149886.

XX
PA (IOWA) UNIV IOWA RES FOUND.

XX
PA (REGC) UNIV CALIFORNIA.

XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX
XX WPI; 2001-234911/24.

XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of

CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.

XX
SQ Sequence 12 AA;

Query Match 86.4%; Score 57; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIK 13
Db 1 RRIIRKIIHIK 12

RESULT 15

AAB70652
ID AAB70652 standard; peptide; 18 AA.

XX
AC AAB70652;

XX
DT 15-MAY-2001 (first entry)

XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:5.

XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;

KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;

KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

XX
OS Ovis aries.

XX
PN WO200112668-A1.

XX
PD 22-FEB-2001.

XX
PF 18-AUG-2000; 2000WO-US22781.

XX
PR 18-AUG-1999; 99US-0149886.

XX
PA (IOWA) UNIV IOWA RES FOUND.

XX
PA (REGC) UNIV CALIFORNIA.

XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX
XX WPI; 2001-234911/24.

XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 10; Page 107; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.

XX
SQ Sequence 18 AA;

Query Match 63.6%; Score 42; DB 22; Length 18;
Best Local Similarity 64.3%; Pred. No. 2.6;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LRRIRKIRK 14
| | | : | | | : | |
Db 3 LrrlgRklahgvkk 16

Search completed: July 12, 2002, 08:00:42
Job time: 449 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:55:05 ; Search time 75.52 seconds
(without alignments)
4.528 Million cell updates/sec

Title: US-09-642-744B-11
Perfect score: 66
Sequence: 1 LRRIIRKIHIIK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backlitest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	38	57.6	457	3	US-09-416-213-2
2	36	54.5	26	1	US-07-725-331-61
3	36	54.5	26	5	PCT-US91-05047-61
4	36	54.5	27	2	US-08-505-486-65
5	36	54.5	27	3	US-08-801-028-65
6	36	54.5	27	3	US-09-340-154-65
7	36	54.5	27	5	PCT-US95-09338-65
8	36	54.5	27	5	PCT-US95-09339-65
9	36	54.5	30	1	US-07-725-331-62
10	36	54.5	30	5	PCT-US91-05047-62
11	36	54.5	36	5	US-07-725-331-63
12	36	54.5	36	5	PCT-US91-05047-63
13	35	53.0	237	3	US-08-388-353-643
14	35	53.0	237	3	US-08-488-551B-643
15	35	53.0	855	1	US-08-022-835-6
16	35	53.0	855	1	US-08-388-809-6
17	35	53.0	855	2	US-08-647-714-6
18	34.5	52.3	475	2	US-08-484-200-2
19	34.5	52.3	475	3	US-08-465-375-2
20	34.5	52.3	475	4	US-08-764-870-10
21	34.5	52.3	475	4	US-08-980-115-10
22	34.5	51.5	669	2	US-07-861-800-2
23	33	50.0	18	5	PCT-US91-05047-16
24	33	50.0	18	5	PCT-US91-05047-16
25	33	50.0	20	1	US-08-786-748A-91
26	33	50.0	20	1	US-08-786-748A-95
27	33	50.0	20	2	US-08-932-682-91

28	33	50.0	20	2	US-08-932-682-95	Sequence 95, Appl
29	33	50.0	21	1	US-08-786-748A-87	Sequence 87, Appl
30	33	50.0	21	2	US-08-932-682-87	Sequence 87, Appl
31	33	50.0	28	1	US-08-786-748A-74	Sequence 74, Appl
32	33	50.0	28	1	US-08-786-748A-77	Sequence 74, Appl
33	33	50.0	28	2	US-08-932-682-74	Sequence 77, Appl
34	33	50.0	28	2	US-08-932-682-77	Sequence 77, Appl
35	33	50.0	76	2	US-08-117-952-754	Sequence 754, App
36	33	50.0	193	2	US-08-850-119-3	Sequence 3, Appl1
37	33	50.0	762	4	US-09-228-986-114	Sequence 114, App
38	33	50.0	3066	4	US-08-952-127-12	Sequence 12, Appl
39	32.5	49.2	18	4	US-08-992-877-62	Sequence 62, Appl
40	32	48.5	15	3	US-08-849-486-10	Sequence 10, Appl
41	32	48.5	18	3	US-09-276-202-7	Sequence 7, Appl1
42	32	48.5	40	1	US-08-179-632-3	Sequence 3, Appl1
43	32	48.5	40	1	US-08-440-174A-3	Sequence 3, Appl1
44	32	48.5	40	5	PCT-US95-00062-3	Sequence 3, Appl1
45	32	48.5	67	3	US-09-120-365-96	Sequence 96, Appl

ALIGNMENTS

RESULT 1
US-09-416-213-2
; Sequence 2, Application US/09416213
; Patent No. 6110690
; GENERAL INFORMATION:
; APPLICANT: Goeddel, David V.
; APPLICANT: Jiang, Yingping
; TITLE OF INVENTION: Suppressors of Death Domains
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/416,213
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 09/035,676
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: T98-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 457 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-416-213-2

Query Match 57.6%; Score 38; DB 3; Length 457;
Best Local Similarity 60.0%; Pred. No. 71;
Matches 6; Conservative 4; Mismatches 0; Gaps 0;
Oy 5 IRRIIRK 14
|:||||:|

DB 380 IKIITHVLEK 389

```
RESULT 2
US-07-725-331-61
; Sequence 61, Application US/07725331
; Patent No. 5294605
; GENERAL INFORMATION:
; APPLICANT: Houghten, Richard
; APPLICANT: Blondelle, Sylvie
; TITLE OF INVENTION: Amphiphilic Peptide Compositions and
; TITLE OF INVENTION: Analogues Thereof
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
; STREET: 180 No. 5294605th Stetson
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/725,331
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/554,422
; FILING DATE: 19-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamsom, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: 421250-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3126165418
; TELEFAX: 3126165460
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: acetylated at N-terminus, may be
; OTHER INFORMATION: a C-terminal amide.
US-07-725-331-61

Query Match 54.5%; Score 36; DB 1; Length 26;
Best Local Similarity 28.6%; Pred. No. 10;
Matches 4; Conservative 9; Mismatches 1; Indels 0; Gaps 0;
```

OY 1 LRIIRKIHIHKK 14
|:|:|:|:|:|:|
Db 1 LKKLKKLKKLKK 14

```
RESULT 3
PCT-US91-05047-61
; Sequence 61, Application PC/TUS9105047
; GENERAL INFORMATION:
; APPLICANT: Houghten, Richard
; APPLICANT: Blondelle, Sylvie
; TITLE OF INVENTION: Amphiphilic Peptide Compositions and
; TITLE OF INVENTION: Analogues Thereof
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
```

```
ADDRESSEE: 6 Milnamow
STREET: 180 North Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/05047
; FILING DATE: 19910717
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/554,422
; FILING DATE: 19-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamsom, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: 421250-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3126165418
; TELEFAX: 3126165460
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: acetylated at N-terminus, may be
; OTHER INFORMATION: a C-terminal amide.
PCT-US91-05047-61
```

OY 1 LRIIRKIHIHKK 14
|:|:|:|:|:|:|
Db 1 LKKLKKLKKLKK 14

```
Query Match 54.5%; Score 36; DB 5; Length 26;
Best Local Similarity 28.6%; Pred. No. 10;
Matches 4; Conservative 9; Mismatches 1; Indels 0; Gaps 0;
```

```
RESULT 4
US-08-505-486-65
; Sequence 65, Application US/08505486
; Patent No. 5955573
; GENERAL INFORMATION:
; APPLICANT: Jesse M. Jaynes
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 Thirteenth Street N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: IBM COMPATIBLE
; OPERATING SYSTEM: DOS
; SOFTWARE: wordperfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,486
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 536
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/279,472
FILING DATE: 22-JUL-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: WALKER, BARBARA W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 2093-117A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ. ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-505-486-65

Query Match 54.5%; Score 36; DB 2; Length 27;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRIIRKIHIKK 14
: || |||: |||
DB 3 KRIARKILKRIKK 15

RESULT 5
US-08-801-028-65
Sequence 65, Application US/08801028
Patent No. 6018102
GENERAL INFORMATION:
APPLICANT: JOAN GABBARINO
APPLICANT: JESSE M. JAYNES
APPLICANT: WILLIAM BELKNAP
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRODUCT
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: STEVEN J. HULTQUIST
ADDRESS: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
STREET: 200 PARK DRIVE, SUITE 210
CITY: RESEARCH TRIANGLE PARK
STATE: NORTH CAROLINA
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
OPERATING SYSTEM: MACINTOSH
SOFTWARE: M.S. WORD 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,028
FILING DATE: 19-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/279,472
FILING DATE: JULY 22, 1994
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-20-94
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-08-94
APPLICATION NUMBER: 08/039,620
FILING DATE: 06-04-93
APPLICATION NUMBER: 08/148,491
FILING DATE: 11-08-93

APPLICATION NUMBER: 08/148,889
FILING DATE: 11-08-93
ATTORNEY/AGENT INFORMATION:
NAME: WASSEMAN, FRANK S.
REGISTRATION NUMBER: 34,273
REFERENCE/DOCKET NUMBER: 4013-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)990-9531
TELEFAX: (919)990-9532
INFORMATION FOR SEQ. ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-801-028-65

Query Match 54.5%; Score 36; DB 3; Length 27;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRIIRKIHIKK 14
: || |||: |||
DB 3 KRIARKILKRIKK 15

RESULT 6
US-09-340-154-65
Sequence 65, Application US/09340154
Patent No. 6084156
GENERAL INFORMATION:
APPLICANT: JESSE M. JAYNES
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street N.W.
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
OPERATING SYSTEM: DOS
SOFTWARE: Wordperfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/340,154
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/505,486
FILING DATE: 21-JUL-1995
APPLICATION NUMBER: U.S. 08/279,472
FILING DATE: 22-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: WALKER, BARBARA W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 2093-117A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ. ID NO: 65:
SEQUENCE CHARACTERISTICS:

LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-09-340-154-65

Query Match 54.5%; Score 36; DB 3; Length 27;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRIIRKIIHIKK 14
:11111:111
Db 3 KRIARKILKRIKK 15

RESULT 7
PCT-US95-09338-65
Sequence 65, Application PC/TUS9509338
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
NUMBER OF SEQUENCES: 98
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09338
FILING DATE: 21-JUL-1994
PRIOR APPLICATION DATA: 08/279,472
APPLICATION NUMBER: 08/279,472
FILING DATE: 22-JUL-1994
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US95-09338-65

Query Match 54.5%; Score 36; DB 5; Length 27;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 RRIIRKIIHIKK 14
:11111:111
Db 3 KRIARKILKRIKK 15
RESULT 8
PCT-US95-09339-65
Sequence 65, Application PC/TUS9509339
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE

TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
NUMBER OF SEQUENCES: 98
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09339
FILING DATE: 21-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/279,472
FILING DATE: 22-JUL-1994
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US95-09339-65

Query Match 54.5%; Score 36; DB 5; Length 27;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRIIRKIIHIKK 14
:11111:111
Db 3 KRIARKILKRIKK 15

RESULT 9
US-07-725-331-62
Sequence 62, Application US/07725331
Patent No. 5294605
GENERAL INFORMATION:
APPLICANT: Houghten, Richard
APPLICANT: Blondelle, Sylvie
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
TITLE OF INVENTION: Analogues Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Sutter, Shore,
ADDRESS: 6 Milnamow
STREET: 180 No. 5294605th Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/725,331
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:

TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
US-07-725-331-62

Query Match 54.5%; Score 36; DB 1; Length 30;
Best Local Similarity 28.6%; Pred. No. 12;
Matches 4; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

OY 1 LRRIRKIIHIKK 14
Db 5 LKKLKKLKLKLLK 18

RESULT 10
PCT-US91-05047-62
Sequence 62, Application PC/TUS9105047
GENERAL INFORMATION:
APPLICANT: Houghten, Richard
APPLICANT: Blondelle, Sylvie
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
STREET: 180 North Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/05047
FILING DATE: 19910717
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamsen, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
PCT-US91-05047-62

Query Match 54.5%; Score 36; DB 5; Length 30;
Best Local Similarity 28.6%; Pred. No. 12;
Matches 4; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

OY 1 LRRIRKIIHIKK 14
Db 5 LKKLKKLKLKLLK 18

RESULT 11
US-07-725-331-63
Sequence 63, Application US/07725331
Patent No. 5294605
GENERAL INFORMATION:
APPLICANT: Houghten, Richard
APPLICANT: Blondelle, Sylvie
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
STREET: 180 No. 5294605th Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/725,331
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamsen, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
US-07-725-331-63

Query Match 54.5%; Score 36; DB 1; Length 36;
Best Local Similarity 28.6%; Pred. No. 14;
Matches 4; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

OY 1 LRRIRKIIHIKK 14
Db 11 LKKLKKLKLKLLK 24

RESULT 12
PCT-US91-05047-63
Sequence 63, Application PC/TUS9105047
GENERAL INFORMATION:
APPLICANT: Houghten, Richard

APPLICANT: Blondelle, Sylvie
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
TITLE OF INVENTION: Analogues Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Suter, Shore,
ADDRESSEE: 6 Milnamow
STREET: 180 North Steelson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/05047
FILING DATE: 19910717
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
OTHER INFORMATION: a C-terminal amide
PCT-US91-05047-63

Query Match 54.5%; Score 36; DB 5; Length 36;
Best Local Similarity 28.6%; Pred. No. 14;
Matches 4; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LRRIRKIHIIKK 14
|:::|:::|:::|
Db 11 LKKLKKLKLKK 24

RESULT 13
US-08-388-353-643
Sequence 643, Application US/08388353
Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Learmont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 643:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-388-353-643

Query Match 53.0%; Score 35; DB 3; Length 237;
Best Local Similarity 50.0%; Pred. No. 1,1e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LRRIRKIHIIKK 14
|:::|:::|:::|
Db 214 LRRIRKIHIIKK 227

RESULT 14
US-08-488-551B-643
Sequence 643, Application US/08488551B
Patent No. 6015661
GENERAL INFORMATION:
APPLICANT: Nicholas J. Deacon
APPLICANT: David Cooper
APPLICANT: Dale A. McPhee
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 841
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,551B
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PM3864 (AU)
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: PM4002 (AU)
FILING DATE: 21-FEB-1994
APPLICATION NUMBER: PM0284 (AU)
FILING DATE: 23-DEC-1994
APPLICATION NUMBER: US 08/388,353
FILING DATE: 14-FEB-1995
APPLICATION NUMBER: PN3021/95
FILING DATE: 17-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FRANK S. DIGIGLIO

REFERENCE/DOCKET NUMBER: 96062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
INFORMATION FOR SEQ ID NO: 643:
SEQUENCE CHARACTERISTICS:
LENGTH: 237 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-551B-643

Query Match 53.0%; Score 35; DB 3; Length 237;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LRRIRKIHIIKK 14
||| |::| |:
Db 214 LRRAYRAILHPRR 227

RESULT 15
US-08-022-835-6
Sequence 6, Application US/08022835
Patent No. 5420030
GENERAL INFORMATION:
APPLICANT: Reitz Jr., Marvin S.
APPLICANT: Franchini, Genoveffa
APPLICANT: Markham, Phillip D.
APPLICANT: Gallo, Robert C.
APPLICANT: Lort, Franco C.
APPLICANT: Popovic, Mikulas
APPLICANT: Garner, Suzanne
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: Eleventh Floor, 1615 L. Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/022,835
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/599,491
FILING DATE: 17-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-022-835-6

Query Match 53.0%; Score 35; DB 1; Length 855;
Best Local Similarity 42.98%; Pred. No. 3.7e+02;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 1 LRRIRKIHIIKK 14
||| |::| |:
Db 832 LQRAVRAILHPRR 845

Search completed: July 12, 2002, 07:55:06
Job time: 113 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:56:50 ; Search time 95.47 Seconds
(without alignments)
14.091 Million cell updates/sec

Title: US-09-642-744B-11
Perfect score: 66
Sequence: 1 LRRIKRIHIKK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	63.6	125	2	A99114
2	42	63.6	152	2	S68411
3	42	63.6	160	2	S68228
4	42	63.6	160	2	S68412
5	41	62.1	41	2	S77768
6	41	62.1	337	2	S46010
7	41	62.1	613	2	F64056
8	40	60.6	370	2	C70464
9	40	60.6	867	2	G69485
10	39	59.1	202	1	H70318
11	39	59.1	210	2	AC1577
12	39	59.1	407	2	T37888
13	39	59.1	483	2	B82160
14	38	57.6	87	2	S00180
15	38	57.6	102	2	AD1420
16	38	57.6	102	2	AE1795
17	38	57.6	272	2	AB3430
18	38	57.6	284	2	A10897
19	38	57.6	623	2	S73462
20	38	57.6	698	2	D64084
21	38	57.6	780	2	AB1801
22	38	57.6	1188	2	C11231
23	37	56.1	334	2	B72301
24	37	56.1	352	2	H86454
25	37	56.1	368	2	C69984
26	37	56.1	569	2	C83952
27	37	56.1	594	2	D70127
28	36	54.5	117	2	T12722
29	36	54.5	234	2	I40822

30	36	54.5	234	2	D97108	DNA-dependent RNA
31	36	54.5	299	2	G72307	hypothetical prote
32	36	54.5	310	2	A33489	hypothetical prote
33	36	54.5	357	2	S21992	envelope protein g
34	36	54.5	370	2	A05034	hypothetical prote
35	36	54.5	447	2	F81678	GTP-binding protei
36	36	54.5	484	2	A70558	hypothetical prote
37	36	54.5	499	2	G84972	Leucyl aminopeptid
38	36	54.5	627	2	C82418	GdDEF family prote
39	36	54.5	672	2	S46276	acetate--CoA ligas
40	36	54.5	760	2	T24521	hypothetical prote
41	36	54.5	780	2	AC1427	transcription regu
42	36	54.5	1014	2	T30840	serine-repeat anti
43	36	54.5	8243	2	T31307	type I fatty acid
44	35	53.0	62	1	R5833F	ribosomal protein
45	35	53.0	98	2	T22503	hypothetical prote

ALIGNMENTS

RESULT 1
A99114
hypothetical protein orf125 [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: A99114
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.;
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MID:11323671
A:Accession: A99114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <DOU>
A:Cross-references: GB:AJ010592; NID:g12580762; PIDN:CAC27080.1; GSPDB:GN00151
C:Genetics:

A:Gene: orf125
A:Map position: 2
A:Genome: nucleomorph
C:Keywords: nucleomorph

Query Match 63.6%; Score 42; DB 2; Length 125;
Best Local Similarity 70.0%; Pred. No. 5.8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 IRKRIHIKK 14
DB 9 IRKRIHLVRK 18

RESULT 2
S68411
cathelin-related protein 2 precursor - sheep (fragment)
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68411
R:Mahoney, M.M.; Lee, A.Y.; Brezinski-Calliguri, D.J.; Huttner, K.M.
FEBS Lett. 377, 519-522, 1995
A:Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobia
A:Reference number: S68411; MID:96140581
A:Accession: S68411
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-152 <MAH>
A:Cross-references: EMBL:X92757
C:Genetics:
A:Gene: SC5-2
C:Superfamily: cathelin; cystatin homology
F:1-21/Domain: signal sequence #status predicted <SIG>
F:14-122/Domain: cystatin homology <CYS>

F:21-123/Domain: propeptide #status predicted <PRO>
F:124-152/Product: cathelin-related protein 2 #status predicted <MAT>

Query Match 63.6%; Score 42; DB 2; Length 152;
Best Local Similarity 64.3%; Pred. No. 7;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LRRIRKIIHIKK 14
|||: ||| | : ||
DB 126 LRRLRKRIAHGVKK 139

RESULT 3

S68228
myeloid antimicrobial peptide 29 precursor - sheep
N:Alternate names: cathelicidin
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68228
R:Bagella, L.; Scocchi, M.; Zanetti, M.
FBS Lett. 376, 225-228, 1995
A:Title: cDNA sequences of three sheep myeloid cathelicidins.
A:Reference number: S68228; MUID:96105386
A:Accession: S68228
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-160 <BAG>
A:Cross-references: EMBL:L46854; NID:g1161248; PIDN:AAA85470.1; PID:g1161249
C:Superfamily: cathelin; cystatin homology
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-130/Domain: cystatin homology <CYS>
F:29-131/Domain: propeptide #status predicted <PRO>
F:132-160/Product: myeloid antimicrobial peptide 29 #status predicted <MAT>

Query Match 63.6%; Score 42; DB 2; Length 160;
Best Local Similarity 64.3%; Pred. No. 7.4;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LRRIRKIIHIKK 14
|||: ||| | : ||
DB 134 LRRLRKRIAHGVKK 147

RESULT 4

S68412
cathelin-related protein 1 precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68412
R:Mahoney, M.M.; Lee, A.Y.; Brezinski-Caliquri, D.J.; Huttner, K.M.
FEBS Lett. 377, 519-522, 1995
A:Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobial F
A:Reference number: S68411; MUID:96140581
A:Accession: S68412
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-160 <MAH>
A:Cross-references: EMBL:X92758
C:Genetics:
A:Gene: SC5-1
C:Superfamily: cathelin; cystatin homology
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-130/Domain: cystatin homology <CYS>
F:30-131/Domain: propeptide #status predicted <PRO>
F:132-160/Product: cathelin-related protein 1 #status predicted <MAT>

Query Match 63.6%; Score 42; DB 2; Length 160;
Best Local Similarity 64.3%; Pred. No. 7.4;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LRRIRKIIHIKK 14
|||: ||| | : ||
DB 134 LRRLRKRIAHGVKK 147

RESULT 5

S77768
hypothetical protein MC037 - Mycoplasma capricolium (fragment)
C:Species: Mycoplasma capricolium
C:Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 21-Jul-2000
C:Accession: S77768; S48590
R:Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, M.L. Microbiol. 16, 955-967, 1995
A:Title: Exploring the Mycoplasma capricolium genome: a minimal cell reveals its phys
A:Reference number: S77739; MUID:96059641
A:Accession: S77768
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-41 <BOB>
A:Cross-references: EMBL:Z33033; NID:g541693; PIDN:CAA83711.1; PID:g4379130
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Genetic code: SGC3

Query Match 62.1%; Score 41; DB 2; Length 41;
Best Local Similarity 57.1%; Pred. No. 3;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LRRIRKIIHIKK 14
| : | : |||||
DB 19 LKERIMQMIIHIKK 32

RESULT 6

S46010
hypothetical protein YBR141c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YBR118
C:Species: Saccharomyces cerevisiae
C:Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 22-Oct-1999
C:Accession: S46010
R:Becam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45995
A:Accession: S46010
A:Molecule type: DNA
A:Residues: 1-337 <BBC>
A:Cross-references: EMBL:Z36010; NID:g536439; PID:g536440; GSPDB:GN00002; MIPS:YBR141
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YBR141c
A:Map position: 2R

Query Match 62.1%; Score 41; DB 2; Length 337;
Best Local Similarity 61.5%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRIIRKIIHIKK 14
||| : | : | : |
DB 30 RRIIRPFHHLINK 42

RESULT 7

F64056
probable ATP-dependent RNA helicase - Haemophilus influenzae (strain Rd Kw20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb-2001
C:Accession: F64056
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kervajave
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: F64056
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-613 <NIGR>
 A:Cross-references: GB:U32709; GB:I42023; NID:g1573190; PIDN:AAC21900.1; PID:g1573195; T
 C:Keywords: ATP; nucleotide binding; P-loop
 F:49-56/Region: nucleotide-binding motif A (P-loop)
 F:151-156/Region: nucleotide-binding motif B
 F:153-158/Region: DEAD motif

Query Match
 Best Local Similarity 62.1%; Score 41; DB 2; Length 613;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14
 ||:| | |:
 Db 353 RRLRNIEHLMK 365

RESULT 8
 C70464
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Feb-2001
 C:Accession: C70464
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666
 A:Accession: C70464
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-370 <AOE>
 A:Cross-references: GB:AE000762; NID:g2984163; PIDN:AAC07703.1; PID:g2984170; GB:AE00065
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: hflX
 C:Superfamily: GTP-binding protein hflX; translation elongation factor Tu homology
 C:Keywords: GTP binding; nucleotide binding; P-loop
 F:197-319/Domain: translation elongation factor Tu homology <EU>
 F:203-210/Region: nucleotide-binding motif A (P-loop)
 F:316-319/Region: GTP-binding NKXD motif

Query Match
 Best Local Similarity 60.6%; Score 40; DB 2; Length 370;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14
 ||:| | |:
 Db 159 KRLIKRIHIKK 171

RESULT 9
 G69485
 DNA-directed RNA polymerase, subunit A' (rpoA1) homolog - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Jun-1999
 C:Accession: G69485
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.F.;
 Gilek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Usterbeck, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
 A:Reference number: A69250; MUID:98049343

A:Accession: G69485
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-867 <KLE>
 A:Cross-references: GB:AE000972; GB:AE000782; NID:g2689295; PIDN:AAB89365.1; PID:g264
 C:Superfamily: Halobacterium DNA-directed RNA polymerase chain A

Query Match
 Best Local Similarity 60.6%; Score 40; DB 2; Length 867;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 LRRIIRKIIHIKK 14
 |||||:| |:
 Db 854 VRLRIYIEGVKK 867

RESULT 10
 H70318
 conserved hypothetical protein aq_202 - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999
 C:Accession: H70318
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666
 A:Accession: H70318
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-202 <AOE>
 A:Cross-references: GB:AE000678; NID:g2982921; PIDN:AAC06551.1; PID:g2982934; GB:AE00
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: aq_202
 C:Superfamily: Methanococcus jannaschii conserved hypothetical protein Mj0226

Query Match
 Best Local Similarity 59.1%; Score 39; DB 1; Length 202;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 3 RRIIRKIIHIKK 13
 |:||:|:
 Db 186 RALRLVHVK 196

RESULT 11
 AC1577
 precorin isomerase homolog lin156 [imported] - Listeria innocua (strain C11p11262)
 C:Species: Listeria innocua
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C:Accession: AC1577
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Esnlt,
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.;
 Ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Roland, J.A.; Voss, H.; Wehla
 A:Title: Comparative genomics of Listeria species
 A:Reference number: AB1077; MUID:21537279; PMID:11679666
 A:Accession: AC1577
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-210 <GIA>
 A:Cross-references: GB:AL592022; PIDN:CAC96387.1; PID:g16413615; GSPDB:GN00178
 A:Experimental source: strain C11p11262
 C:Genetics:
 A:Gene: lin156
 C:Superfamily: Methanobacterium thermoautotrophicum precorin isomerase

```

Query Match      59.1%; Score 39; DB 2; Length 210;
Best Local Similarity 60.0%;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Oy 4 IIRKIIHIIK 13
      .:.:|:|:|:|:|
Db 62 VIOKIIHVLK 71

```

RESULT 12
 T37888
 hypothetical protein SPAC1805.03c - fission yeast (*Schizosaccharomyces pombe*)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T37888
 R:Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL data library, September 1999
 A:Reference number: Z21752
 A:Accession: T37888
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-407 <RIE>
 A:Cross-references: EMBL:AL117390; PIDN:CAB56844.1; GSPDB:GN000066; SPDB:SPAC1805.03c
 A:Experimental source: strain 972h; cosmid cl805
 C:Genetics:
 A:Gene: SPDB:SPAC1805.03c
 A:Map position: 1
 A:Introns: 32/2; 148/2; 160/3

```

Query Match Similarity      59.1%; Score 39; DB 2; Length 407;
Best Local Similarity      50.0%; Pred. No. 56;
Matches      6; Conservative      4; Mismatches      2; Indels      0; Gaps      0;

OY      3 RIIRKIHIIRK 14
      1 :|||1:::1
Db      6 RFLRKIAHLVOK 17

RESULT 13
B82160
hypothetical protein VC1762 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C$Species: Vibrio cholerae
C$Date: 18-Aug-2000 #sequence-revision 20-Aug-2000 #text_change 02-Feb-2001
C$Accession: B82160
R$Heidelberg, J.F.; Eelsen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
charidson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R.
l, R.R.; Metlalnios, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A$Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A$Reference number: A82035; MUID:20406833
A$Accession: B82160
A$status: preliminary
A$molecule type: DNA
A$Residues: 1-483 <HEI>
A$Cross-references: GB:AE004253; GB:AE003852; NID:9956278; PIDN:AAF94911.1; GSPDB:GN001
A$Experimental source: serogroup O1; strain N16961; biotype El Tor
C$Genetics:
A$Gene: VC1762
A$Map position: 1

```

Query March	59.1%	Score 39;	DB 2;	Length 483;
Best Local Similarity	53.8%	Pred. No. 66;		
Matches	7;	Conservative	3;	Mismatches
			3;	Indels
			0;	Gaps
OY	2	RRIRKRIIHIKK	14	
		::	1::	
db	91	RRLLKRIEIVAK	103	

RESULT 14
S00180

```

C: spermated protein SL - smaller spotted catshark
C: Species: Scyliorhinus canicula (smaller spotted catshark, smaller spotted dogfish)
C: Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 23-Feb-1997
C: Accession: S00180
R: Chauviere, M.; Martinge, A.; Briand, G.; Sautiere, P.; Chevallerier, P.
Eur. J. Biochem. 169, 105-111, 1987
A: Title: Nuclear basic protein transltion during sperm differentiation: amino acid se
A: Reference number: S00180; MUID:88055030
A: Accession: S00180
A: Molecule type: protein
A: Residues: 1-87 <CH>
C: Superfamily: sperm histone
C: Keywords: DNA binding; nucleus; phosphoprotein; sperm
F: 3.55/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match          57.6%  Score 38:  DB 2:  Length 87:
Matches 6:  Similarity 46.2%  Pred. No. 19:
        6:  Conservative 5:  Mismatches 2:  Indels 0:  Gaps 0:

```

RESULT 15
AD1420
PTS cellobiose-specific enzyme IIA homolog [mo2765 (imported)] - *Listeria monocytogenes*
C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AD1420
R:Glaser, P., Frangoul, L., Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecher, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihi, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schlutener, T.; Simoes, N.; Tiersitz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11579669
A:Accession: AD1420
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-102 <GUA>
A:Cross-references: GB:NC_003210; PIDN:CAD00978.1; PID:g16412265; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: Imo2765
C:Superfamily: phosphotransferase system lactose-specific enzyme II, factor III

Query Match	57.6%	Score 38	DB 2	Length 102
Best Local Similarity	35.7%	Pred. No. 23		
Matches	5	Conservative	5	Mismatches
			4	Indels
			0	Gaps
0y	1	LRRIIRKIIRHIHK	14	
		::: ::		
Db	87	LRVLETFVHVYKK	100	

Search completed: July 12, 2002, 07:56:52
Job time: 219 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:04:37 ; Search time 46.15 Seconds
(without alignments)
11.746 Million cell updates/sec

Title: US-09-642-744B-11
Perfect score: 66
Sequence: 1 LRRIRKIHIIKK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query	Match Length	ID	Description
1	42	63.6	160	1	SC51_SHEEP
2	42	63.6	160	1	SC52_SHEEP
3	41	62.1	337	1	YB21_YEAST
4	41	62.1	613	1	DEAD_HAEIN
5	38	57.6	87	1	SSS1_SCYCA
6	38	57.6	457	1	BAG4_HUMAN
7	38	57.6	623	1	Y014_MYCPN
8	38	57.6	670	1	REP_HAEIN
9	37	56.1	368	1	YSFB_BACSU
10	37	56.1	594	1	SYA_BORBU
11	36	54.5	370	1	MATK_MARPO
12	36	54.5	499	1	AMPA_BUCAL
13	36	54.5	590	1	SGT1_SCHPO
14	36	54.5	672	1	ACSA_PHYBL
15	35	53.0	62	1	RL30_BACST
16	35	53.0	164	1	RL21_ARATH
17	35	53.0	180	1	CAS2_RABIT
18	35	53.0	189	1	TBP_AERPE
19	35	53.0	212	1	ALKR_HAEIN
20	35	53.0	364	1	LEU3_AQUAE
21	35	53.0	420	1	PR11_HUMAN
22	35	53.0	437	1	ERG_ARATH
23	35	53.0	438	1	MCRB_METFE
24	35	53.0	529	1	TIMK_METFE
25	35	53.0	529	1	TIM_FCOLI
26	35	53.0	529	1	TIM_SALPO
27	35	53.0	577	1	SYP_HELPY
28	35	53.0	626	1	ACSA_NEUCR
29	35	53.0	661	1	ACSA_COPCI
30	35	53.0	805	1	TF2_AOUAE
31	35	53.0	821	1	MCW6_HUMAN
32	35	53.0	847	1	ENV_HV1W2
33	35	53.0	856	1	ENV_HV1W1

34	34.5	52.3	475	1	PPAT_CRIGR
35	34.5	52.3	505	1	PPAT_MOUSE
36	34.5	52.3	505	1	PBAT_RAT
37	34	51.5	67	1	RL30_RHEMA
38	34	51.5	130	1	CW36_YEAST
39	34	51.5	208	1	IL6_BOVIN
40	34	51.5	262	1	MURI_BUCAL
41	34	51.5	280	1	Y174_TREPA
42	34	51.5	299	1	MODP_RHIME
43	34	51.5	400	1	ENPL_MESAU
44	34	51.5	410	1	CH60_EHRCA
45	34	51.5	417	1	PR11_MOUSE

ALIGNMENTS

RESULT ID	1	SC51_SHEEP	STANDARD:	PRT:	160 AA.
AC	P49928:	SC51_SHEEP			
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-NOV-1997	(Rel. 35, Last annotation update)			
DE	Cathelin-related peptide SC5 precursor 1 (Antibacterial peptide SMAP-29).				
DE	29) (Myeloid antibacterial peptide SMAP-29).				
OS	Ovis aries (Sheep).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Caprinae; Ovis.				
OX	NCBI_TaxID=9940;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RP	TISSUE=Bone marrow;				
RC	MEDLINE=96140581; PubMed=8549789;				
RX	MAHONEY M.M., Lee A.Y., Brezinski-Calliquiri D.J., Hutner K.M.;				
RA	"Molecular analysis of the sheep cathelin family reveals a novel				
RT	antimicrobial peptide.";				
RL	FEBS Lett. 377:519-522(1995).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RP	TISSUE=Liver;				
RC	Hutner K.M., Mahoney M.M.;				
RL	Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.				
CC	-1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.				
CC	-1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.				
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: X92757; CA63412.1; -				
DR	EMBL: U60600; AAB49715.1; -				
DR	InterPro: IPR001894; Cathelicidin.				
DR	Pfam: PF00666; Cathelicidins; 1.				
DR	Prodom: PD001838; Cathelicidin; 1.				
DR	PROSITE: PS00946; CATHELICIDINS_1; 1.				
DR	PROSITE: PS00947; CATHELICIDINS_2; 1.				
KW	Antibiotic; Signal.				
FT	SIGNAL	1	29	POTENTIAL.	
FT	PROPEP	30	131	BY SIMILARITY.	
FT	CHAIN	132	160	CATHELIN-RELATED PEPTIDE SC5.	
FT	MOD_RES	30	30	PYROLIDONE CARBOXYLIC ACID	
FT				(BY SIMILARITY).	
FT	DISULFID	86	97	BY SIMILARITY.	
FT	DISULFID	108	125	BY SIMILARITY.	
SO	SEQUENCE	160 AA;	17786 MW;	BD9B3859CA32C249 CRC64;	

Query Match 63.6% Score 42; DB 1; Length 160;

Best Local Similarity 64.3%; Pred. NO. 1.5;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LRRIRKIIHIKK 14

Db 134 LRRIGRKIIAHGVKK 147

```

RESULT 2
SC52_SHEEP STANDARD; PRT; 160 AA.
ID SC52_SHEEP
AC P49929;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cathelin-related peptide SC5 precursor 2 (Antibacterial peptide SMAP-29) (Myeloid antibacterial peptide SMAP-29).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=96140581; PubMed=8549789;
RA Mahoney M.M., Lee A.Y., Brezinski-Calisuri D.J., Huttner K.M.;
RT "Molecular analysis of the sheep cathelin family reveals a novel
   antimicrobial peptide."
RL FEBS Lett. 377:519-522(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=96105386; PubMed=7498547;
RA Bagella L., Scocchi M., Zanetti M.;
RT "cDNA sequences of three sheep myeloid cathelicidins."
RL FEBS Lett. 376:225-228(1995).
CC -1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X92758; CAA63413.1; -
DR EMBL; LA6854; AA85470.1; -
DR InterPro; IPR001894; Cathelicidin.
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00666; Cathelicidins; 1.
DR ProDom; PD001838; Cathelicidin; 1.
DR SMART; SM00043; CY; 1.
DR PROSITE; PS00946; CATHELICIDINS_1; 1.
DR PROSITE; PS00947; CATHELICIDINS_2; 1.
KW Antibiotic; Signal.
FT SIGNAL 1 29
FT PROPEP 30 131
FT CHAIN 132 160
FT MOD_RES 30 30
FT DISULFID 86 97
FT DISULFID 108 125
FT CONFLICT 28 28
FT SEQUENCE 160 AA; 17742 MW; 4FB98A09355BB51F CRC64;

```

Query Match 63.6% Score 42; DB 1; Length 160;
Best Local Similarity 64.3%; Pred. NO. 1.5;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LRRIRKIIHIKK 14

Db 134 LRRIGRKIIAHGVKK 147

```

RESULT 3
YB21_YEAST STANDARD; PRT; 337 AA.
ID YB21_YEAST
AC P38278;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 38.5 kDa protein in IRA1-MAK5 intergenic region.
GN YBR141C OR YBR1118.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=95274325; PubMed=7754712;
RA Zagulski M., Becam A.-M., Grzybowska E., Lacroute F., Migdalski A.,
RA Slonimski P.P., Sokolowska B., Herbert C.J.;
RT "The sequence of 12.5 kb from the right arm of chromosome II predicts
   a new N-terminal sequence for the IRA1 protein and reveals two new
   genes, one of which is a DEAD-box helicase."
RL Yeast 10:1227-1234(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z36010; CAA85099.1; -
DR EMBL; X78937; CAA35358.1; -
DR PIR; S46010; S46010.
DR SGD; S0000345; YBR141C.
KW Hypothetical protein.
SO SEQUENCE 337 AA; 38539 MW; CB0039B18FADF3BE CRC64;

```

Query Match 62.1% Score 41; DB 1; Length 337;
Best Local Similarity 61.5%; Pred. NO. 4.9;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14

Db 30 RRIIRFPHLINK 42

```

RESULT 4
DEAD_HAEIN STANDARD; PRT; 613 AA.
ID DEAD_HAEIN
AC P44586;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cold-shock DEAD-box protein A homolog (ATP-dependent RNA helicase dead
   homolog).
GN DEAD OR CSDA OR H10231.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;

```

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kellaway A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McEwen K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.,
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.,"
RL Science 269:496-512(1995).
CC -1- FUNCTION: HAS A HELIX-DESTABILIZING ACTIVITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32709; AAC21900.1; -;
DR HSP; Q58083; IHV8.
DR TIGR: H10231; -;
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000629; DEAD_ATP_helicase.
DR InterPro: IPR001650; Helicase-C.
DR Pfam: PF00270; DEAD.1.
DR Pfam: PF00271; Helicase-C.1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELICC; 1.
DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
DR Hydroxylase: Helicase; ATP-binding; RNA-binding;
KM Transcription regulation; Complete proteome.
FT NE-BIND 49 56
FT SITE 155 158 DEAD BOX.
SQ SEQUENCE 613 AA; 69705 MW; 1B826CBDEB1704DF CRC64;

Query Match 62.1%; Score 41; DB 1; Length 613;
Best Local Similarity 53.8%; Pred. No. 9;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIHIKK 14
DB 353 RRLRNIEHLMKK 365

RESULT 5
SSSI_SCYCA STANDARD; PRT; 87 AA.
AC P13275;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 01-JAN-1990 (Rel. 13, Last annotation update)
OS Spermatid-specific protein S1.
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeoidea; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.
ON NCBI_TaxID=7830;
RX [1]
RP SEQUENCE:
RA MEDLINE=88055030; PubMed=3678229;
RA Chauviere M., Martineau A., Briand G., Sautiere P., Chevallier P.;
RT "Nuclear basic protein transition during sperm differentiation. Amino
RT acid sequence of a spermatid-specific protein from the dog-fish
RT Scyllorhinus caniculus.";
RL Eur. J. Biochem. 169:105-111(1987).
CC -1- FUNCTION: INVOLVED IN NUCLEAR BASIC PROTEIN TRANSITION: HISTONES

CC ARE REPLACED BY SPERMATID SPECIFIC PROTEINS WHICH ARE THEMSELVES
CC REPLACED BY PROTAMINES IN LATE SPERMATIDS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: WITH PROTEIN S2.
DR PIR: S00180; S00180.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Nuclear protein.
SQ SEQUENCE 87 AA; 11179 MW; 38A0ED7A8256910 CRC64;

Query Match 57.6%; Score 38; DB 1; Length 87;
Best Local Similarity 46.2%; Pred. No. 4.2;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIHIKK 14
DB 42 RRRRKIVHLKRR 54

RESULT 6
BAG4_HUMAN STANDARD; PRT; 457 AA.
AC O95429; O95818;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE BAG-family molecular chaperone regulator-4 (Silencer of death
DE domains).
GN BAG4 OR SODD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9901615; PubMed=9873016;
RX Takayama S., Xie Z., Reed J.C.;
RT "An evolutionarily conserved family of Hsp70/Hsc70 molecular
RT chaperone regulators.";
RT J. Biol. Chem. 274:781-786(1999).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=99115917; PubMed=9915703;
RX Jiang Y., Woronicz J.D., Liu W., Goeddel D.V.;
RT "Prevention of constitutive TNF receptor 1 signaling by silencer of
RT death domains.";
RL Science 283:543-546(1999).
RN [3]
RN ERRATUM.
RP Jiang Y., Woronicz J.D., Liu W., Goeddel D.V.;
RL Science 283:1852-1852(1999).
CC -1- FUNCTION: INHIBITS THE CHAPERONE ACTIVITY OF HSP70/HSC70 BY
CC PROMOTING SUBSTRATE RELEASE.
CC -1- SIMILARITY: BINDS TO THE ATPASE DOMAIN OF HSP70/HSC CHAPERONES.
CC -1- SIMILARITY: CONTAINS 1 BAG DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF095194; AAD16123.2; -;
DR EMBL: AF111116; AAD05226.1; -;
DR MIM: 603884; -;
DR InterPro: IPR003103; BAG.
DR Pfam: PF02179; BAG.1.
DR SMART: SM00264; BAG; 1.
KW Chaperone.
FT DOMAIN 379 456 BAG.
SQ SEQUENCE 457 AA; 49594 MW; B89D59E8118684A3 CRC64;

Query Match 57.6%; Score 38; DB 1; Length 457;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 5 IRIKIHIKK 14
1:1111111111
DB 380 IKIIVLEK 389

RESULT 7
Y014_MYCPN STANDARD; PRT; 623 AA.
ID Y014_MYCPN
AC P75095; 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ABC transporter ATP-binding protein MG014 homolog (D12_orf623)
GN MPN018 OR MP136.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=9710585; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSHA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000015; AAB95784.1; -.
DR HSSP: P13569; INBD.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmem.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane_1.
DR Pfam: PF00005; ABC_tran; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW Hypothetical protein; ATP-binding; Transport; Transmembrane;
KW Complete proteome.
FT TRANSMEM 27 47 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 157 177 POTENTIAL.
FT TRANSMEM 180 200 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 307 327 POTENTIAL.
FT NP_BIND 400 407 ATP (POTENTIAL).
SQ SEQUENCE 623 AA; 68928 MW; ABD0481D8AEF736B CRC64;

Query Match 57.6%; Score 38; DB 1; Length 623;
Best Local Similarity 70.0%; Pred. No. 31;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 RRIIRKIHI 11
1:1111111111

DB 117 RKIFRKIMHI 126

RESULT 8
REP_HAEIN STANDARD; PRT; 670 AA.
ID REP_HAEIN
AC P44804;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ATP-dependent DNA helicase rep (EC 3.6.1.-).
GN REP OR H10649.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Pine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Guelm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd";
RL Science 269:496-512(1995).
CC -1- FUNCTION: REP HELICASE IS A SINGLE-STRANDED DNA-DEPENDENT
CC ATPASE INVOLVED IN DNA REPLICATION. IT CAN INITIATE UNWINDING
CC AT A NICK IN THE DNA. IT BINDS TO THE SINGLE-STRANDED DNA AND
CC ACTS IN A PROGRESSIVE FASHION ALONG THE DNA IN THE 3' TO 5'
CC DIRECTION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32748; AAC22309.1; -.
DR HSSP: P09980; UUA.
DR TIGR: H10649; -.
DR InterPro: IPR000212; UVRD-helicase.
DR Pfam: PF00580; UVRD-helicase; 1.
KW Hydrolyase; Helicase; DNA replication; ATP-binding; DNA-binding;
KW Complete proteome.
FT NP_BIND 22 29 ATP (POTENTIAL).
SQ SEQUENCE 670 AA; 77745 MW; 802A131251D4E24 CRC64;

Query Match 57.6%; Score 38; DB 1; Length 670;
Best Local Similarity 61.5%; Pred. No. 33;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRIIRKIHIKK 14
1:1111111111
DB 30 RVIIRKIIVLEK 42

RESULT 9
YSFB_BACSU STANDARD; PRT; 368 AA.
ID YSFB_BACSU
AC P94536;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 42.6 kDa protein in sspi-csta intergenic region.
GN ysfB.
OS Bacillus subtilis.
OC Bacteria: Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97124191; PubMed=8969504;
RA Wpat A., Carter N., Biggell C.S., Guy J.B., Piper K., Sanders J.,
RA Emerson P.T., Haywood C.R.;
RT "The dnaB-phA (256 degrees-240 degrees) region of the Bacillus
RT subtilis chromosome containing genes responsible for stress responses,
RT the utilization of plant cell walls and primary metabolism.";
RL Microbiology 142:3067-3078(1996).
CC -! SIMILARITY: BELONGS TO THE CDAR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 275208; CAA99600.1; -
CC DR EMBL: 299118; CAB14827.1; -
CC DR Subtilist: BG12319; ysfB.
CC KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 368 AA; 42611 MW; 37E1C096AF6EF03C CRC64;

Query Match 56.1%; Score 37; DB 1; Length 368;
Best Local Similarity 66.7%; Pred. NO. 27;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 RIIRKIHI 11
   |||:||||
Db 214 RLRIKILHI 222

RESULT 10
SYA_BORBU STANDARD; PRT; 594 AA.
ID 051238;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alanyl-tRNA synthetase (BC 6.1.1.7) (Alanine--tRNA ligase) (Alars).
DE A1AS OR B80220.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kellavag A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utermbeck T., Matthey L., McDonald L., Atliach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi.";
RX Nature 390:580-586(1997).
RL Nature 390:580-586(1997).
CC -! CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC -! SUBCELLULAR LOCATION: Cytoplasmic.
CC -! SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001132; AAC6604.1; -
CC DR TIGR: B80220; -
CC DR InterPro: IPR002106; AA_tRNA_ligase_II.
CC DR InterPro: IPR002318; tRNA-synt_2c.
CC DR Pfam: PF01411; tRNA-synt_2c; 2.
CC DR PRINTS: PRO0960; TRANASYTHALA.
CC DR PROSITE: PS00179; AA_tRNA_LIGASE_II_1; FALSE_NEG.
CC DR PROSITE: PS00339; AA_tRNA_LIGASE_II_2; FALSE_NEG.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC complete proteome.
SQ SEQUENCE 594 AA; 67773 MW; 2337B306FD02349A CRC64;

Query Match 56.1%; Score 37; DB 1; Length 594;
Best Local Similarity 57.1%; Pred. NO. 44;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 LRRIIRKIHIHIKK 14
   |||:||||
Db 307 LRRIIRSIYVAKK 320

RESULT 11
MATK_MARPO STANDARD; PRT; 370 AA.
ID MATK_MARPO
AC P12174;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Probable intron maturase (Maturase K).
DN MATK OR YCF14.
OS Marchantia polymorpha (Liverwort).
OC Chloroplast.
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
OC Marchantiales; Marchantiaceae; Marchantia.
OX NCBI_TaxID=3197;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S.,
RA Ohysono K., Shiki Y., Takeuchi M., Chang Z., Aota S., Inokuchi H.,
RA Ozeki H.;
RT "Chloroplast gene organization deduced from complete sequence of
RT liverwort Marchantia polymorpha chloroplast DNA.";
RL Nature 322:572-574(1986).
CC -! FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC INTRONS.
CC -! SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED
CC BY MITOCHONDRIAL INTRONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X04465; CAA28076.1; -
CC DR PIR: A05034; A05034.
CC DR PIR: S01589; S01589.
CC DR Mendel: 5294; MARPO.ycf14.1.
CC DR InterPro: IPR000442; intron_maturase2.
CC DR InterPro: IPR002866; MATK_N.

```

DR Pfam; PF01348; Intron_maturase2; 1.
 DR Pfam; PF01824; MatK_N; 1.
 KM Chloroplast; mRNA processing.
 SQ SEQUENCE 370 AA; 45407 MW; 0917812AD6DC0495 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 370;
 Best Local Similarity 70.0%; Pred. No. 41;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 5 RKIHIKK 14
 :|:| |:-|
 Db 109 IKKIHIKK 118

RESULT 12
 ID AMPA_BUCAI STANDARD; PRT; 499 AA.
 AC P57448;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE CytoSOL aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase) (LAP)

DE (Leuyl aminopeptidase).

GN PEPA OR BU367

OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum

OS symbiotic bacterium).

OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.

OX NCBI_TaxID=118099;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-TOKYO 1998;

RX MEDLINE-20445173; PubMed-10993077;

RA Shigenobu S., Matanabe H., Hattori M., Sakaki Y., Ishikawa H.;

RT "Genome sequence of the endocellular bacterial symbiont of aphids

RT Buchnera sp. APS.";

RL Nature 407:81-86(2000).

CC -!- FUNCTION: PRESUMABLY INVOLVED IN THE PROCESSING AND REGULAR

CC TURNOVER OF INTRACELLULAR PROTEINS. CATALYZES THE REMOVAL OF

CC UNSUBSTITUTED N-TERMINAL AMINO ACIDS FROM VARIOUS PEPTIDES (BY

CC SIMILARITY).

CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-

CC xbb-, in which Xaa is preferably Leu, but may be other amino acids

CC including Pro although not Arg or Lys, and xbb may be Pro.

CC -!- COFACTOR: MANGANESE (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17; ALSO KNOWN AS THE

CC CYTOSOL AMINOPEPTIDASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AP001119; BAB13071.1; -

DR InterPro; IPR000819; Peptidase_M17.

DR Pfam; PF00883; Peptidase_M17.1.

DR Pfam; PF02789; Peptidase_M17.N.1.

DR PRINTS; PR00481; LAMNOPTDASE.

DR PROSITE; PS00631; CYTOSOL_AP; 1.

KM Hydrolyase; Aminopeptidase; Manganese; Complete proteome.

FT METAL 267 272 MANGANESE 2 (BY SIMILARITY).

FT METAL 272 272 MANGANESE 1 AND 2 (BY SIMILARITY).

FT METAL 290 290 MANGANESE 2 (BY SIMILARITY).

FT METAL 349 349 MANGANESE 1 (BY SIMILARITY).

FT METAL 351 351 MANGANESE 1 AND 2 (BY SIMILARITY).

FT ACT_SITE 279 279 POTENTIAL.

FT ACT_SITE 353 353 POTENTIAL.

SQ SEQUENCE 499 AA; 55268 MW; DA7DB26BBE24A957 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 499;
 Best Local Similarity 46.2%; Pred. No. 55;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 RKIHIKK 14
 :|:| |:-|
 Db 90 KRIKKIHIKK 102

RESULT 13
 ID SGT1_SCHPO STANDARD; PRT; 590 AA.
 AC Q90S49;

DT 01-MAR-2002 (Rel. 41, Created)

DT 01-MAR-2002 (Rel. 41, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE SGT1 protein homolog Cl002.10c.

GN SPAC1002.10c.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-972;

RA Cadieu E., Lelaure V., Galbert F., McDougall R.C., Rajandream M.A.,

RA Barrell B.G.;

RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE SGT1 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AL136078; CAB65610.1; -

DR Hypothetical protein.

KM SEQUENCE 590 AA; 68156 MW; 1BD07A283EB0294C CRC64;

SQ SEQUENCE 590 AA; 68156 MW; 1BD07A283EB0294C CRC64;

OY 6 RKIHIKK 14
 :|:| |:-|
 Db 215 RIVHVLQK 223

RESULT 14
 ID ACSA_PHYBL STANDARD; PRT; 672 AA.
 AC Q01576;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Acetyl-coenzyme A synthetase (EC 6.2.1.1) (acetate--CoA ligase) (Acyl-

DE activating enzyme).

GN FACA.

OS Phycomyces blakesleeana.

OC Eukaryota; Fungi; zygomycota; zygomycetes; Mucorales; Mucoraceae;

OC Phycomyces

OX NCBI_TaxID=4837;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NRRL 1555;

RX MEDLINE-94335878; PubMed-7914670;

RA Garre V., Murillo F.J., Torres-Martinez S.;

RT "Isolation of the facA (acetyl-CoA synthetase) gene of *Phycomyces*
 RT *blakesleeanus*.";
 RL Mol. Gen. Genet. 244:278-286(1994).
 CC -1- CATALYTIC ACTIVITY: ATP + acetate + CoA = AMP + diphosphate +
 CC acetyl-CoA.
 CC -1- INDUCTION: BY ACETATE.
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M94729; AA53586.1; .
 DR HSSP: P08659; ILCI.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Ligase.
 SO SEQUENCE 672 AA; 74775 MW; DA6BCBA35252034E CRC64;

Query Match 54.5%; Score 36; DB 1; Length 672;
 Best Local Similarity 60.0%; Pred. No. 74;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 LRRIRKIIH 10
 :||:||||:
 Db 627 MRRIRKIVN 636

RESULT 15
 RL30_BACST STANDARD; PRT; 62 AA.
 ID P02431;
 AC 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last annotation update)
 DE 50S ribosomal protein L30.
 GN RPLM.
 OS *Bacillus stearothermophilus*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC *Bacillus/Staphylococcus* group; *Geobacillus*.
 OX NCBI_TaxID=142;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=8411493; PubMed=6363400;
 RA Kimura M.;
 RT "Proteins of the *Bacillus stearothermophilus* ribosome. The amino acid
 RT sequences of proteins S5 and L30.";
 RL J. Biol. Chem. 259:1051-1055(1984).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=87016925; PubMed=3463963;
 RA Wilson K.S., Appelt K., Badger J., Tanaka I., White S.W.;
 RT "Crystal structure of a prokaryotic ribosomal protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:7251-7255(1986).
 CC -1- SIMILARITY: BELONGS TO THE L30P FAMILY OF RIBOSOMAL PROTEINS.
 CC PIR: A02827; R5BS3F.
 DR HSSP: P74909; IBXY.
 DR InterPro: IPR000517; Ribosomal_L30.
 DR Pfam: PF00327; Ribosomal_L30; 1.
 DR PROSITE: PS00634; RIBOSOMAL_L30; 1.
 KW Ribosomal protein.
 SO SEQUENCE 62 AA; 7053 MW; AE96D2CDF5C56253 CRC64;

Query Match 53.0%; Score 35; DB 1; Length 62;
 Best Local Similarity 38.5%; Pred. No. 9.9;

Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 OY 1 LRRIRKIIH 13
 :|:|:|:|:
 Db 43 IRGMIRVAHLVK 55

Search completed: July 12, 2002, 08:04:38
 Job time: 650 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:03:46 ; Search time 174.7 Seconds
(without alignments)
13.863 Million cell updates/sec

Title: US-09-642-744B-11

Perfect score: 66

Sequence: 1 LRRIRKIIHIIRK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mmc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rv1virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	63.6	125	10	Q9AVZ1
2	42	63.6	500	3	Q9P6R0
3	41	62.1	41	2	Q48971
4	41	62.1	217	1	Q977R2
5	41	62.1	414	10	Q947U2
6	40	60.6	370	16	Q67743
7	40	60.6	867	17	Q28391
8	39	59.1	202	16	Q66580
9	39	59.1	210	16	Q92CL5
10	39	59.1	407	3	Q9UTR1
11	39	59.1	483	16	Q9KR81
12	39	59.1	838	15	Q9DVL4
13	39	59.1	903	5	Q95MA7
14	38	57.6	102	16	Q926T2
15	38	57.6	147	11	Q91VT5
16	38	57.6	415	10	Q9SP04

17	38	57.6	457	11	Q9CWC2	Q9CWC2 mus musculu
18	38	57.6	491	4	Q9H8X2	Q9H8X2 homo sapien
19	38	57.6	780	16	Q926T7	Q926T7 listeria in
20	38	57.6	841	15	Q41556	Q41556 human immun
21	38	57.6	871	15	Q91U27	Q91U27 human immun
22	38	57.6	954	5	Q9GQ15	Q9GQ15 giardia lam
23	38	57.6	969	5	Q9U022	Q9U022 giardia lam
24	38	57.6	1188	17	Q57849	Q57849 pyrococcus
25	37	56.1	295	3	Q13635	Q13635 schizosacch
26	37	56.1	334	16	Q9X0D8	Q9X0D8 thermotoga
27	37	56.1	349	10	Q9SRT5	Q9SRT5 arabidopsis
28	37	56.1	352	10	Q9MAP0	Q9MAP0 arabidopsis
29	37	56.1	385	2	P94599	P94599 bacillus th
30	37	56.1	407	10	Q9SFW0	Q9SFW0 arabidopsis
31	37	56.1	460	5	Q9VU03	Q9VU03 drosophila
32	37	56.1	569	16	Q9KA71	Q9KA71 bacillus ha
33	37	56.1	578	5	Q95XS3	Q95XS3 caenorhabdi
34	37	56.1	694	10	Q94G68	Q94G68 oryza sativ
35	37	56.1	857	15	Q73288	Q73288 human immun
36	37	56.1	858	15	Q73290	Q73290 human immun
37	37	56.1	858	15	Q73291	Q73291 human immun
38	37	56.1	858	15	Q73292	Q73292 human immun
39	37	56.1	859	15	Q71260	Q71260 human immun
40	37	56.1	861	15	Q91U26	Q91U26 human immun
41	36	54.5	56	15	Q56190	Q56190 human immun
42	36	54.5	56	15	Q56192	Q56192 human immun
43	36	54.5	117	9	Q80196	Q80196 methanobact
44	36	54.5	158	5	Q904W9	Q904W9 drosophila
45	36	54.5	158	5	Q9V3S6	Q9V3S6 drosophila

ALIGNMENTS

RESULT 1

ID Q9AVZ1 PRELIMINARY: PRT: 125 AA.

AC Q9AVZ1;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 15.2 KDA PROTEIN.

OS Guillardia theta (Cryptomonas phi).

OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.

OX NCBI_TaxID=55529;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20087226; PubMed=10618395;

RA Zauner S., Fraunholz M., Mastl J., Penny S.L., Beaton M., Cavalier-Smith T., Maier U., Douglas S.;

RT "Chloroplast protein and centosomal genes, a tRNA intron, and odd telomeres in an unusually compact eukaryotic genome, the cryptomonad nucleomorph.";

RT Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21223349; PubMed=11323671;

RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L., Wu X., Reith M., Cavalier-Smith T., Maier U.;

RT "The highly reduced genome of an enslaved algal nucleus.";

RL Nature 410:1091-1096(2001).

DR EMBL: AJ010592; CAC27080.1; -

KW Hypothetical protein.

SQ SEQUENCE 125 AA; 15207 MW; AAD53BB5593997A1 CRC64;

Query Match 63.6%; Score 42; DB 10; Length 125;

Best Local Similarity 70.0%; Pred. No. 10;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 5 IRKIIHIIRK 14

Db 9 IRKIIHLVRK 18

```
RESULT 2
Q9P6R0 ID Q9P6R0 PRELIMINARY; PRT: 500 AA.
AC Q9P6R0;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE TRANSCRIPTION FACTOR IIB 70 KDA SUBUNIT.
GN SPIC137.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=972H-.
RC Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL554632; CAB89885.1; -.
DR InterPro: IPR004366; Cyclicin.
DR InterPro: IPR000812; TE11B.
DR Pfam: PF00382; transcript_fac2; 2.
DR PRINTS: PR00685; TIFACTOR1B.
DR SMART: SM00385; CYCLIN; 2.
SQ SEQUENCE 500 AA; 56749 MW; 6CF1FDF9ADA2B1B2 CRC64;
```

```
Query Match 63.6%; Score 42; DB 3; Length 500;
Best Local Similarity 41.7%; Pred. No. 37;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 RRIIRKIHIHK 13
Db 247 RRSVREVVHVVK 258
```

```
RESULT 3
Q48971 ID Q48971 PRELIMINARY; PRT: 41 AA.
AC Q48971;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE SIMILAR TO DCTP DEAMINASES (FRAGMENT).
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Entomoplasmataceae.
OX NCBI_TaxID=2095;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 27343(KID);
RC MEBLINE=96059641; PubMed=7476192;
RA Bork P., Ouzounis C., Casari G., Schneider R., Sander C., Dolan M.,
RA Gilbert W., Gillevert P.M.;
RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals
its physiology."
RL Mol. Microbiol. 16:955-967(1995).
DR EMBL: Z33033; CAAB3711.1; -.
FT NON TER 1 41
FT NON TER 1 41
SQ SEQUENCE 41 AA; 5021 MW; 8BBB3709DA1EB90B CRC64;
```

```
Query Match 62.1%; Score 41; DB 2; Length 41;
Best Local Similarity 57.1%; Pred. No. 5.4;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 LRIIRKIHIHK 14
Db 19 LKEFIMOMIHIHK 32
```

```
RESULT 4
Q977R2 ID Q977R2 PRELIMINARY; PRT: 217 AA.
AC Q977R2;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE NAD(P)+-FLAVIN OXIDOREDUCTASE.
OS uncultured crenarchaeote 487.
OC Archaea; Crenarchaeota; environmental samples;
OC marine archaeal group 1.
OX NCBI_TaxID=44557;
RN [1]
RP SEQUENCE FROM N.A.
RA Beja O., Koonin E.V., Aravind L., Taylor L.T., Seitz H., Stein J.L.,
RA Bensen D.C., Feldman R.A., Swanson R.V., Delong E.F.;
RT "Comparative genomic analysis of coexisting archaeal genetic variants
in an Antarctic marine microbial assemblage."
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: U40238; AAK66808.1; -.
SQ SEQUENCE 217 AA; 24624 MW; E22C83B6D4957E67 CRC64;
```

```
Query Match 62.1%; Score 41; DB 1; Length 217;
Best Local Similarity 42.9%; Pred. No. 25;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 LRIIRKIHIHK 14
Db 202 LRQKELIHVVK 215
```

```
RESULT 5
Q947U2 ID Q947U2 PRELIMINARY; PRT: 414 AA.
AC Q947U2;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE PUTATIVE TRANSLATION INITIATION FACTOR 2B BETA SUBUNIT.
GN OSJNB40068A07.14.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eurnatoidaeae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NIPPONBARE.
RA de la Bastide M., Spiegel L., Kirchoff K., Preston R., King L.,
RA Vil M.D., Baker J., Zutavern T., Santos L., Miller B., Kuit K.,
RA Cunnius D.M., Bell M., Ballja V., Shan R., Bahret A., Dix S.,
RA Yang C., O'Shaughnessy A., Palmer L., Dechia N., McCombie W.R.;
RT "Genomic Sequence for Oryza sativa, Nipponbare strain, clone
OSJNB40068A07, from Chromosome 10, complete sequence."
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC091734; AAK98757.1; -.
KW Initiation factor.
SQ SEQUENCE 414 AA; 44646 MW; BA54863017C7EF0C CRC64;
```

```
Query Match 62.1%; Score 41; DB 10; Length 414;
Best Local Similarity 54.5%; Pred. No. 46;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 IIRKIHIHK 14
Db 76 IVRVLIHIKE 86
```

```
RESULT 6
```


RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
 RA Madeno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordstie G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remiel B., Rose M., Schueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of *Listeria species.*";
 RL Science 294:849-852(2001).
 DR EMBL: AL596167; CAC96387.1; -.
 DR ListList; LIN01156; -.
 DR Complete proteome.
 SQ SEQUENCE 210 AA; 23107 MW; 205A7B01966F913A CRC64;

Query Match 59.1%; Score 39; DB 16; Length 210;
 Best Local Similarity 60.0%; Pred. No. 54;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 4 RIIRKIH1K 13
 DB 62 VIQKIHVK 71

RESULT 10
 O9UTH1 PRELIMINARY; PRT; 407 AA.
 AC O9UTH1:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 47.1 KDA PROTEIN.
 GN SPAC1805.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972H-;
 RA Rieger M., McDougall R.C., Rajandream M.A., Barrell B.G.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL117390; CAB55844.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 407 AA; 47107 MW; C1077AE50E3EE28 CRC64;

Query Match 59.1%; Score 39; DB 3; Length 407;
 Best Local Similarity 50.0%; Pred. No. 99;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 RIIRKIH1K 14
 DB 6 RLRIKIH1VK 17

RESULT 11
 O9KR81 PRELIMINARY; PRT; 483 AA.
 AC O9KR81:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN VCI1762.
 GN VCI1762.
 OS *Vibrio cholerae*.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Ginn M.L.,
 RA Dodson R.D., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,

RA Gill S.R., Nelson K.E., Read T.D., Tetteijn H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi T., Sellers P.,
 RA McDonald L., Ueberback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004253; AAF94911.1; -.
 DR TIGR: VCI1762; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 483 AA; 55076 MW; 34D93FF2A8F1D03C CRC64;

Query Match 59.1%; Score 39; DB 16; Length 483;
 Best Local Similarity 53.8%; Pred. No. 1.2e+02;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 RIIRKIH1K 14
 DB 91 RLRIKIH1VK 103

RESULT 12
 O9DVL4 PRELIMINARY; PRT; 838 AA.
 AC O9DVL4:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ENVELOPE PROTEIN.
 GN ENV OR GP160.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=97DC.KMST120;
 RX MEDLINE=20499072; PubMed=11044094;
 RA Vidal N., Peeters M., Mulanga-Kabeya C., Nzilambi N., Robertson D.,
 RA Ilunga W., Sema H., Tshimanga K., Bongo B., Delaporte E.;
 RT "Unprecedented degree of human immunodeficiency virus type 1 (HIV-1)
 RT group M genetic diversity in the Democratic Republic of Congo suggests
 RT that the HIV-1 pandemic originated in Central Africa."
 RL J. Virol. 74:10498-10507(2000).
 DR EMBL: AJ401039; CAC15047.1; -.
 DR InterPro: IPR000328; Env.GP41.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 DR Pfam: PF00517; GP41; 1.
 KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
 SQ SEQUENCE 838 AA; 95125 MW; 59083EA564F06165 CRC64;

Query Match 59.1%; Score 39; DB 15; Length 838;
 Best Local Similarity 42.9%; Pred. No. 1.9e+02;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 LRRIKIH1K 14
 DB 815 LQRYVRAIH1PRR 828

RESULT 13
 O95W47 PRELIMINARY; PRT; 903 AA.
 AC O95W47:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CIRCADIAN CLOCK PROTEIN PERIOD.
 GN PER.
 OS *Bulla gouldiana* (California bubble).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia;
 OC Cephalaspiidae; Bullidae; Bulla.
 OX NCBI_TaxID=114738;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Constance C.M.;
 RT "Cloning and analysis of clock genes in the marine mollusc, Bulla
 RT Gouldiana."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF353619; AAK97374.1; -
 SQ SEQUENCE 903 AA; 101146 MW; 22BD87376BC5324A CRC64;

Query Match 59.1%; Score 39; DB 5; Length 903;
 Best Local Similarity 50.0%; Pred. No. 2.le+02;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 LRIRIRKIHIIKK 14
 Db 424 LRIRIRKIHIIKK 437

RESULT 14
 O926Y2 PRELIMINARY; PRT; 102 AA.
 AC O926Y2;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 GN LIN2908.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Cherouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Ertian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkut G.,
 RA Madueno E., Maitounam A., Mala Vicente J., Ng E., Nedjari H.,
 RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of Listeria species."
 RL Science 294:849-852(2001).
 DR EMBL: AL596174; CAC98133.1; -
 DR Listlist: LIN02908; -
 KW Complete proteome.
 SQ SEQUENCE 102 AA; 11476 MW; 0561109775B8A334 CRC64;

Query Match 57.6%; Score 38; DB 16; Length 102;
 Best Local Similarity 35.7%; Pred. No. 40;
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 LRIRIRKIHIIKK 14
 Db 87 LRIRIRKIHIIKK 100

RESULT 15
 O91VT5 PRELIMINARY; PRT; 147 AA.
 AC O91VT5;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE SIMILAR TO RIKEN CDNA 2410112115 GENE (FRAGMENT).
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BREAST TUMOR;
 RA Strausberg R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC009102; AA09102.1; -
 FT NON_TER
 SQ SEQUENCE 147 AA; 16518 MW; E9F9640957200DD3 CRC64;

Query Match 57.6%; Score 38; DB 11; Length 147;
 Best Local Similarity 60.0%; Pred. No. 57;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 5 IRIRIRKIHIIKK 14
 Db 70 IRIRIRKIHIIKK 79

Search completed: July 12, 2002, 08:03:47
 Job time: 633 sec


```
XX Claim 1; Page 103; 137pp; English.
PS
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
SQ Sequence 28 AA:

Query Match 100.0%; Score 140; DB 22; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLRRLGRKIAHGKYGKPTVLRIRIA 28
   |||||||
Db 1 rglrrlgrkiahygkkygptvlririia 28

RESULT 2
AAB70649 standard; peptide: 29 AA.
XX
AC AAB70649;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:2.
XX
KM Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
PA (REGC ) UNIV CALIFORNIA.
XX
PI Track BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
DR WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 10; Page 107; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
```

```
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
SQ Sequence 29 AA:

Query Match 100.0%; Score 140; DB 22; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.2e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGLRRLGRKIAHGKYGKPTVLRIRIA 28
   |||||||
Db 1 rglrrlgrkiahygkkygptvlririia 28

RESULT 3
AAB70651 standard; peptide: 20 AA.
XX
AC AAB70651;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:4.
XX
KM Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
PA (REGC ) UNIV CALIFORNIA.
XX
PI Track BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
DR WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 10; Page 107; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
SQ Sequence 20 AA:

Query Match 72.9%; Score 102; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LGRKIAHGKYGKPTVLRIRI 25
   |||||||
```

Db	1	lgrkiahgvkkygptvlrli	20
RESULT	4		
AAAB70675			
ID	AAAB70675	standard; peptide: 29 AA.	
AC	AAAB70675;		
XX			
DT	15-MAY-2001	(first entry)	
XX			
DE	Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:28.		
XX			
KW	Ovine: SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;		
KM	bactericidal; antibiotic; antiviral; microbial growth inhibitor;		
KW	proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;		
KM	Burkholderia cepacia; Alcaligenes; Xanthomonas.		
XX			
OS	Ovis aries.		
XX			
PN	WO200112668-A1.		
XX			
PD	22-FEB-2001.		
XX			
PF	18-AUG-2000; 2000WO-US22781.		
XX			
PR	18-AUG-1999; 99US-0149886.		
XX			
PA	(IOWA) UNIV IOWA RES FOUND.		
XX	(REGC) UNIV CALIFORNIA.		
PI	Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;		
XX			
DR	WPI; 2001-234911/24.		
PT	New antimicrobial peptides useful as antibiotics for inhibiting growth		
PT	and proliferation of microbes, and for treating microbial infections -		
XX			
XX	Claim 1; Page 103; 137pp; English.		
PS			
XX			
CC	AAAB70648 to AAAB70675 represent antimicrobial peptides (I), of which		
CC	AAAB70648 to AAAB70664, AAAB70674 and AAAB70675 are derived from the ovine		
CC	SMAP 29 cathelicidin family peptide, and AAAB70665 to AAAB70673 are		
CC	derived from the lupine RCAP 18 cathelicidin family peptide. (I) have		
CC	antibiotic, antimicrobial and antiviral activities, and can be used as		
CC	microbial growth and proliferation inhibitors and in gene therapy. (I)		
CC	are useful for inhibiting microbial growth in an environment capable of		
CC	sustaining such growth, for inhibiting microbial growth or strain in a		
CC	host, and inhibiting the growth of drug-resistant microbial strains such		
CC	as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and		
CC	Xanthomonas.		
XX			
XX	Sequence 29 AA:		
SO			
Query Match	70.0%;	Score 98;	DB 22;
Best Local Similarity	70.4%;	Pred. No. 7.3e-08;	Length 29;
Matches 19;	Conservative 4;	Mismatches 4;	Indels 0;
QY	1	RGRLRLGKRIAHGVKGYPTVLRIRI	27
Db	1	:	
1	knrlrlrkriahikkygptvlrli	27	
RESULT	5		
AAAB70652			
ID	AAAB70652	standard; peptide: 18 AA.	
AC	AAAB70652;		
XX			
DT	15-MAY-2001	(first entry)	
XX			
DE	Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:5.		

XX	Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW	bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW	proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW	Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX	
OS	Ovis aries.
XX	
FN	WO200112668-A1.
PD	22-FEB-2001.
XX	
PF	18-AUG-2000; 2000MO-US22781.
XX	
PR	18-AUG-1999; 99US-0149886.
XX	
PA	(IOWA) UNIV IOWA RES FOUND.
PA	(REGC) UNIV CALIFORNIA.
XX	
PI	Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX	
DR	WPI; 2001-234911/24.
PT	New antimicrobial peptides useful as antibiotics for inhibiting growth
PT	and proliferation of microbes, and for treating microbial infections -
XX	
PS	Claim 10; Page 107; 137p; English.
XX	
CC	AA870648 to AAB70675 represent antimicrobial peptides (1), of which
CC	AA870648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC	SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC	derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC	antibiotic, antimicrobial and antiviral activities, and can be used as
CC	microbial growth and proliferation inhibitors and in gene therapy. (1)
CC	are useful for inhibiting microbial growth in an environment capable of
CC	sustaining such growth, for inhibiting microbial growth or strain in a
CC	host, and inhibiting the growth of drug-resistant microbial strains such
CC	as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC	Xanthomonas.
XX	
SQ	Sequence 18 AA:
	Query Match 67.1%, Score 94; DB 22; Length 18;
	Best Local Similarity 100.0%; Pred. No. 1.7e-07;
	Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
OY	1 KGLRLGKRIAGVKKYG 18
DB	1 rglrlgkrlahgvkkyg 18
RESULT 6	
AAAB70650	
ID	AAAB70650 standard; peptide: 21 AA.
XX	
AC	AAAB70650;
XX	
DT	15-MAY-2001 (first entry)
XX	
DE	Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:3.
XX	
KW	Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW	bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW	proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW	Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX	
OS	Ovis aries.
XX	
PN	WO200112668-A1.
XX	
PD	22-FEB-2001.
XX	

PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
XX (REGC) UNIV CALIFORNIA.
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
DR
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
XX
PS Claim 10; Page 107; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
SQ Sequence 21 AA:

Query Match 65.7%; Score 92; DB 22; Length 21;
Best Local Similarity 95.0%; Pred. No. 4.1e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 9 KIAHGKRYGPTVLRIRIA 28
: ||||| ||||| |||||
Db 1 KIAHGKRYGPTVLRIRIA 20

RESULT 7
AAB70648
ID AAB70648 standard; peptide; 18 AA.
XX
AC AAB70648;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:1.
XX
XX Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
XX (REGC) UNIV CALIFORNIA.
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth

PT and proliferation of microbes, and for treating microbial infections
XX
PS Claim 1; Page 103; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
SQ Sequence 18 AA:

Query Match 40.7%; Score 57; DB 22; Length 18;
Best Local Similarity 61.1%; Pred. No. 0.069;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 RGLRLGRKRIAHGVKYG 18
: |||: ||| | : ||||
Db 1 KRLRLGRKRIAHGVKYG 18

RESULT 8
AAB70655
ID AAB70655 standard; peptide; 18 AA.
XX
AC AAB70655;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:8.
XX
XX Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
XX (REGC) UNIV CALIFORNIA.
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
XX
PS Claim 1; Page 103; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a

CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.

XX Sequence 18 AA;

Query Match 40.7%; Score 57; DB 22; Length 18;
 Best Local Similarity 61.1%; Pred. No. 0.069;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGLRLGRKIAHGKVKYK 18
 : ||| ||| | : |||
 Db 1 knlrrlrkrlhikkyg 18

RESULT 9

AA75211
 ID AAY75211 standard; Protein; 449 AA.

XX AAY75211;

DT 21-MAR-2000 (first entry)

DE *Neisseria gonorrhoeae* ORF 603 protein sequence SEQ ID NO:1896.

XX *Neisseria meningitidis*; *Neisseria gonorrhoeae*; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
 KW antibacterial; gene therapy.

XX *Neisseria gonorrhoeae*.

XX MO9957280-A2.

PD 11-NOV-1999.

PF 30-APR-1999; 99WO-US09346.

PR 01-MAY-1998; 98US-0083758.

PR 31-JUL-1998; 98US-0094869.

PR 02-SEP-1998; 98US-0098994.

PR 09-OCT-1998; 98US-0103749.

PR 09-OCT-1998; 98US-0103794.

PR 09-OCT-1998; 98US-0103796.

PR 25-FEB-1999; 99US-0121528.

XX (CHIR) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;

PI Petersen J, Piza M, Rappoli R, Ratti G, Scalato E, Scarselli M;

PI Tettein H, Venter JC;

DR MPI: 2000-062150/05.

XX N-PSDB; AA253973.

XX Novel *Neisseria* polypeptides predicted to be useful antigens for

PT vaccines and diagnostics

XX Claim 2; Page 955; 1453pp; English.

XX AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941

CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides

CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent

CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.

XX Sequence 449 AA;

Query Match 40.0%; Score 56; DB 21; Length 449;
 Best Local Similarity 50.0%; Pred. No. 3.5;
 Matches 10; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 3 LRLGRKIAHGKVKYKPTVL 22
 : : ||| ||| : ||| : |||
 Db 133 lkaigrriahgkkyhesvl 152

RESULT 10

AA70654
 ID AAB70654 standard; peptide; 18 AA.

XX AAB70654;

DT 15-MAY-2001 (first entry)

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:7.

XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
 KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;

XX *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

XX Ovis aries.

XX WO200112668-A1.

PD 22-FEB-2001.

PF 18-AUG-2000; 2000WO-US22781.

PR 18-AUG-1999; 99US-0149886.

PR (IOWA) UNIV IOWA RES FOUND.

PA (REGC) UNIV CALIFORNIA.

XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX MPI: 2001-234911/24.

XX New antimicrobial peptides useful as antibiotics for inhibiting growth

PT and proliferation of microbes, and for treating microbial infections

XX Claim 1; Page 103; 137pp; English.

XX AAB70648 to AAB70675 represent antimicrobial peptides (1), of which

CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine

CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are

CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have

CC antibiotic, antimicrobial and antiviral activities, and can be used as

CC microbial growth and proliferation inhibitors and in gene therapy. (1)

CC are useful for inhibiting microbial growth in an environment capable of

CC sustaining such growth, for inhibiting microbial growth or strain in a

CC host, and inhibiting the growth of drug-resistant microbial strains such

CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and

CC *Xanthomonas*.

XX Sequence 18 AA;

Query Match 39.3%; Score 55; DB 22; Length 18;
 Best Local Similarity 55.6%; Pred. No. 0.14;
 Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RGLRLGRKIAHGKVKYK 18
 : : ||| ||| | : |||

[illegible]

KM Escherichia coli; identification; proliferation; microorganism;
KM antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;
XX bacterial growth inhibition.
OS
XX Escherichia coli.
PN WO200148209-A2.
XX
XX 05-JUL-2001.
PD
XX 19-DEC-2000; 2000MO-US34419.
PF
XX 23-DEC-1999; 99US-0173005.
PR
XX (ELIT-) ELITRA PHARM INC.
PA
XX Forsyth RA, Ohlsen KL, Zyskind JW;
PI WPI: 2001-457376/49.
XX N-PSDB; AAH81365.
DR
XX Novel nucleic acids encoding proteins required for Escherichia coli
PT proliferation, useful for screening for antimicrobial agents -
PS Claim 19; Page 485-486; 596pp; English.
XX
XX The present invention describes a purified or isolated nucleic acid
CC sequence (I) consisting essentially of one of the 93 nucleotide sequences
CC given in AAH81202 to AAH81294, where expression of the nucleic acid in a
CC microorganism is capable of inhibiting proliferation of a microorganism.
CC (I) have antibacterial and antibiotic activities, and can be used in
CC gene therapy. Expression of (I) in a microorganism inhibits proliferation
CC of the microorganism, and the manufactured antibiotic is useful for
CC reducing the activity or level of a gene product required for
CC proliferation of a microorganism in a subject, specifically humans. The
CC nucleic acids that inhibit bacterial growth or proliferation can be used
CC as antisense therapeutics for killing bacteria. In addition to
CC therapeutic applications, the nucleic acid sequences complementary to
CC sequences required for proliferation can be used as diagnostic tools.
CC For example, nucleic acid probes complementary to proliferation-required
CC sequences that are specific for particular species of microorganisms can
CC be used as probes to identify particular microorganism species in
CC clinical specimens. AAH81295 to AAH81487 encode the Escherichia coli
CC proteins given in AAG984239 to AAG98431, and AAH81488 to AAH81491
CC represent oligonucleotides, which are used in the exemplification of the
CC present invention.
XX
XX
SQ Sequence 367 AA;

Query Match 38.6%; Score 54; DB 22; Length 367;
Best Local Similarity 59.1%; Pred. No. 5.7;
Matches 13; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

OY 7 GRKIAHGKVKYGPVLRIRIA 28
||| | | | | | : | | | | |
Db 320 grdiaghvrynsdalt--ria 339

RESULT 14
ID AAW37784 standard; Protein: 375 AA.
XX
XX AAW37784;
AC
XX
XX 28-JUL-1998 (first entry)
DT
XX
XX Plasmid pPF1 structural protein lacZ.Nterm-proB74.
DE
XX
XX L-proline-4-hydroxylase; food additive; transformant; anti-inflammatory;
KM trans-4-hydroxy-L-proline; carbapenem antibiotic; proline biosynthesis;
KW pharmaceutical intermediate; proa. ss.
XX

OS Synthetic.
XX
XX EP826773-A2.
PN
XX
XX 04-MAR-1998.
PD
XX
XX 03-SEP-1997; 97EP-0115289.
PF
XX
XX 03-SEP-1996; 96JP-0232724.
PR
XX
XX (KIOW) KYOWA HAKKO KOGYO KK.
PA
XX
XX Maruyama A, Mori H, Motoyama H, Ozaki A, Shibasaki T;
PI WPI: 1998-161081/15.
XX N-PSDB; AAV19021.
DR
XX
XX Transformant for trans-4-hydroxy-L-proline production - comprising
PT proline-over-producing host transformed with L-proline 4-hydroxylase
PT gene
PS Example 3; Page 17-18; 21pp; English.
XX
XX This sequence represents a synthetic structural protein found in
CC plasmid pPF1 and is composed of a gamma-glutamyl kinase including the
CC N-terminal, 8 amino acid residues (lacZ.Nterm) and a gene, proa, under the
CC beta-galactosidase including the N-terminal and a gene, proa, under the
CC control of plac. This plasmid is used in a method involving the
CC construction of an L-proline-4-hydroxylase high expression plasmid to
CC produce transformant comprising a proline-overproducing host containing a
CC recombinant DNA construct. The transformant can be cultured to produce
CC trans-4-hydroxy-L-proline, which is a pharmaceutical intermediate (e.g.
CC for producing carbapenem antibiotics and the anti-inflammatory agent
CC N-acetyl-4-hydroxyproline) and a food additive.
XX
XX
SQ Sequence 375 AA;

Query Match 38.6%; Score 54; DB 19; Length 375;
Best Local Similarity 59.1%; Pred. No. 5.8;
Matches 13; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

OY 7 GRKIAHGKVKYGPVLRIRIA 28
||| | | | | | : | | | | |
Db 328 grdiaghvrynsdalt--ria 347

RESULT 15
ID AAY75212 standard; Protein: 450 AA.
XX
XX AAY75212;
AC
XX
XX 21-MAR-2000 (first entry)
DT
XX
XX Neisseria meningitidis ORF 603 protein sequence SEQ ID NO:1898.
DE
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KM antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
XX
XX Neisseria meningitidis.
OS
XX
XX WO9957280-A2.
PN
XX
XX 11-NOV-1999.
PD
XX
XX 30-APR-1999; 99MO-US09346.
PF
XX
XX 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
XX 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.

PR 09-OCT-1998; 98US-0103749.
 PR 09-OCT-1998; 98US-0103794.
 PR 09-OCT-1998; 98US-0103796.
 PR 25-FEB-1999; 99US-0121528.

XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX

PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,
 PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M,
 PI Tettelin H, Venter JC;

XX
 DR WPI: 2000-062150/05.
 DR N-PSDB; AAZ53974.

PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX

PS Claim 2; Page 956; 1453pp; English.

XX
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ54253 to AAZ55941
 CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.

XX
 SQ Sequence 450 AA;

Query Match 37.9%; Score 53; DB 21; Length 450;

Best Local Similarity 45.0%; Pred. No. 10;
 Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRRIGRKRIAHGKRYGPTVL 22
 ::||::||::||::||
 Db 134 ikaighriahggekysesvl 153

Search completed: July 12, 2002, 08:00:45
 Job time: 452 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:55:09 ; Search time 75.52 Seconds
(without alignments)
9.056 Million cell updates/sec

Title: US-09-642-744B-27
Perfect score: 140
Sequence: 1 RGLRLGRKRIANGVKKKYPTLIRIRIA 28

Scoring table: BLOSUM62
Gap0 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	35.0	307	3	US-08-605-284B-19
2	49	35.0	1820	2	US-07-998-289B-8
3	49	35.0	2100	3	US-08-808-793-23
4	49	35.0	2100	3	US-08-772-512A-19
5	49	35.0	2104	2	US-08-808-793-4
6	49	35.0	2104	3	US-08-772-512A-4
7	49	35.0	2105	2	US-08-808-793-3
8	49	35.0	2105	3	US-08-772-512A-3
9	48	34.3	501	3	US-08-906-791-2
10	48	34.3	501	4	US-09-111-730-1
11	47	33.6	34	3	US-08-782-997A-20
12	47	33.6	35	1	US-08-039-557-3
13	47	33.6	35	3	US-08-782-997A-19
14	47	33.6	35	3	US-08-782-997A-25
15	47	33.6	35	4	US-08-749-066A-2
16	47	33.6	61	2	US-08-385-590A-9
17	47	33.6	61	3	US-09-021-520-9
18	47	33.6	309	3	US-08-605-284B-4
19	47	33.6	309	3	US-08-605-284B-5
20	47	33.6	309	3	US-08-605-284B-6
21	47	33.6	1011	3	US-08-836-325-2
22	47	33.6	1984	3	US-08-836-325-10
23	47	33.6	1989	3	US-08-836-325-12
24	46	32.9	35	3	US-08-782-997A-24
25	46	32.9	37	4	US-08-749-066A-1
26	46	32.9	38	2	US-08-385-590A-12
27	46	32.9	38	2	US-08-723-306-24

28	46	32.9	38	3	US-09-021-520-12	Sequence 12, Appl
29	46	32.9	38	5	PCT-US96-10041-24	Sequence 24, Appl
30	46	32.9	62	2	US-08-385-590A-16	Sequence 16, Appl
31	46	32.9	62	3	US-09-021-520-16	Sequence 16, Appl
32	46	32.9	68	2	US-08-385-590A-20	Sequence 20, Appl
33	46	32.9	68	3	US-09-021-520-20	Sequence 20, Appl
34	46	32.9	310	3	US-08-605-284B-10	Sequence 10, Appl
35	46	32.9	310	3	US-08-605-284B-13	Sequence 13, Appl
36	46	32.9	310	3	US-08-605-284B-14	Sequence 14, Appl
37	46	32.9	310	3	US-08-605-284B-15	Sequence 15, Appl
38	46	32.9	2005	3	US-08-836-325-7	Sequence 7, Appl
39	45	32.1	35	3	US-08-782-997A-23	Sequence 23, Appl
40	45	32.1	56	4	US-09-269-299A-1	Sequence 1, Appl
41	45	32.1	232	3	US-09-024-020B-6	Sequence 6, Appl
42	45	32.1	232	4	US-09-425-043-6	Sequence 6, Appl
43	45	32.1	311	3	US-08-605-284B-18	Sequence 18, Appl
44	45	32.1	413	2	US-08-808-793-25	Sequence 25, Appl
45	45	32.1	452	3	US-07-998-289B-6	Sequence 6, Appl

ALIGNMENTS

```
RESULT 1
US-08-605-284B-19
: Sequence 19, Application US/08605284B
: Patent No. 6060271
:
: GENERAL INFORMATION:
: APPLICANT: WALEWSKI, JOSE L.
: APPLICANT: RECIO-PINTO, ESPERANZA
: TITLE OF INVENTION: VOLTAGE GATED SODIUM CHANNELS FROM
: TITLE OF INVENTION: HUMAN PERIPHERAL NERVE
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NIXON, HARGRAVE, DEVYANS & DOYLE LLP
: STREET: CLINTON SQUARE, P.O. BOX 1051
: CITY: ROCHESTER
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 14603
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/605,284B
: FILING DATE: 09-FEB-1996
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: BRAMAN, SUSAN J.
: REGISTRATION NUMBER: 34,103
: TELEPHONE/DOCKET NUMBER: 19603/800 (CRF D-1705)
: TELEPHONE: 716-263-1636
: TELEFAX: 716-263-1600
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 307 amino acids
: TYPE: amino acid
: STRANDEDNESS: not relevant
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-605-284B-19
:
: Query Match 35.0% Score 49; DB 3; Length 307;
: Best Local Similarity 40.0%; Pred. No. 5.2;
: Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1;
:
: Oy 6 LGRKAHGVKKY--GPTVLIRIRIA 28
: Db 139 LGVLSDIIKKYFSPILRLRVRA 163
```

```

RESULT 2
US-07-998-289B-8
; Sequence 8, Application US/07998289B
; Patent No. 6027876
; GENERAL INFORMATION:
; APPLICANT: Black, Bruce C
; APPLICANT: Taylor, Marlin
; APPLICANT: Heckel, David G
; TITLE OF INVENTION: Method for Monitoring Pesticide
; TITLE OF INVENTION: Resistance
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10022
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/998, 289B
; FILING DATE: 30-DEC-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Robinson, Joseph R
; REGISTRATION NUMBER: 33,448
; REFERENCE/DOCKET NUMBER: 0646/OA939
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1820 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-07-998-289B-8

Query Match 35.0%; Score 49; DB 3; Length 1820;
Best Local Similarity 40.0%; Pred. No. 39;
Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1;

CY 6 LGRKIAHGKYKT-GPTVLRIRIA 28
II : : : II : : : : :
Db 1651 LGLVLSDIRKYEFSPTLLRVVRA 1675

RESULT 3
US-08-808-793-23
; Sequence 23, Application US/08808793
; Patent No. 5858713
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603

```

```

1      COMPUTER READABLE FORM:
2      MEDIUM TYPE: Floppy disk
3      COMPUTER: IBM PC compatible
4      OPERATING SYSTEM: PC-DOS/MS-DOS
5      SOFTWARE: Patentin Release #1.0, Version #1.30
6      CURRENT APPLICATION DATA:
7      APPLICATION NUMBER: US/08/808,793
8      FILING DATE:
9      CLASSIFICATION: 435
10     PRIOR APPLICATION DATA:
11     APPLICATION NUMBER: US 60/034,361
12     FILING DATE: 24-DEC-1996
13     CLASSIFICATION: 435
14     PRIOR APPLICATION DATA:
15     APPLICATION NUMBER: US 60/012,649
16     FILING DATE: 01-MAR-1996
17     CLASSIFICATION: 435
18     ATTORNEY/AGENT INFORMATION:
19     NAME: Branan, Susan J.
20     REGISTRATION NUMBER: 34,103
21     REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)
22     TELECOMMUNICATION INFORMATION:
23     TELEPHONE: 716-263-1636
24     TELEFAX: 716-263-1600
25     INFORMATION FOR SEQ ID NO: 23:
26     SEQUENCE CHARACTERISTICS:
27     LENGTH: 2100 amino acids
28     TYPE: amino acid
29     STRANDEDNESS: not relevant
30     TOPOLOGY: linear
31     MOLECULE TYPE: protein
32     US-08-808-793-23
33
34     Query Match          35.0%; Score 49; DB 2; Length 2100;
35     Best Local Similarity 40.0%; Pred. No. 46;
36     Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1:
37
38     QY      6 LGRKIAHGKKY--GPTVIRIRIA 28
39             ||::||| ||::||::||
40     Db      1664 LGLVLSDIIEKRYFSPTLIRVRA 1688
41
42     RESULT      4
43     US-08-772-512A-19
44     ; Sequence 19, Application US/08772512A
45     ; Patent No. 6022705
46     ;
47     ; GENERAL INFORMATION:
48     ; APPLICANT: Soderlund, David M.
49     ; APPLICANT: Knipfle, Douglas C.
50     ; APPLICANT: Ingles, Patricia J.
51     ; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM
52     ; TITLE OF INVENTION: INSECTICIDE-SUSCEPTIBLE AND INSECTICIDE-RESISTANT HOUSE
53     ; TITLE OF INVENTION: FILES
54     ; NUMBER OF SEQUENCES: 19
55     ; CORRESPONDENCE ADDRESS:
56     ; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
57     ; STREET: P.O. Box 1051, Clinton Square
58     ; CITY: Rochester
59     ; STATE: New York
60     ; COUNTRY: USA
61     ; ZIP: 14603
62     ;
63     ; COMPUTER READABLE FORM:
64     ; MEDIUM TYPE: Floppy disk
65     ; COMPUTER: IBM PC compatible
66     ; OPERATING SYSTEM: PC-DOS/MS-DOS
67     ; SOFTWARE: Patentin Release #1.0, Version #1.30
68     ; CURRENT APPLICATION DATA:
69     ; APPLICATION NUMBER: US/08/772,512A
70     ; FILING DATE:
71     ; CLASSIFICATION: 435
72     ; PRIOR APPLICATION DATA:
73     ; APPLICATION NUMBER: US 08/608,618
74     ;

```

```

;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 2104 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: not relevant
;

```

RESULT 7
US-08-808-793-3

```
; Sequence 3, Application US/08808793
; Patent No. 5858713
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,793
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/034,361
; FILING DATE: 24-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,649
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Braham, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2105 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-808-793-3

Query Match          35.0%; Score 49; DB 2; Length 2105;
Best Local Similarity 40.0%; Pred. No. 46;
Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1;

QY      6 LGKRIAGVKKY--GPTVLRIRIA 28
      ||:::|||||:|:|:|:|:|
Db      1683 LGLVLSDIIEKYFVSPTLRVRA 1707

RESULT      8
US-08-772-512A-3
; Sequence 3, Application US/08772512A
; Patent No. 6022705
; GENERAL INFORMATION:
; APPLICANT: Soderlund, David M.
; APPLICANT: Knipple, Douglas C.
; APPLICANT: Ingles, Patricia J.
; TITLE OF INVENTION: INSECT SODIUM CHANNELS FROM
; TITLE OF INVENTION: INSECTICIDE-SUSCEPTIBLE AND INSECTICIDE-RESISTANT HOUSE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP
; STREET: P.O. Box 1051, Clinton Square
```

```
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,512A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,618
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Braham, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 19603/601(CRPD-1657)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1636
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2105 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-772-512A-3

Query Match          35.0%; Score 49; DB 3; Length 2105;
Best Local Similarity 40.0%; Pred. No. 46;
Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1;

QY      6 LGKRIAGVKKY--GPTVLRIRIA 28
      ||:::|||||:|:|:|:|:|
Db      1683 LGLVLSDIIEKYFVSPTLRVRA 1707

RESULT      9
US-08-906-791-2
; Sequence 2, Application US/08906791
; Patent No. 6096876
; GENERAL INFORMATION:
; APPLICANT: St-Arnaud, Rene
; APPLICANT: Glorieux, Francis H.
; TITLE OF INVENTION: 1-alpha-Hydroxylase Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Towers, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,791
; FILING DATE: 06-AUG-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Toole, Edward M.
; REGISTRATION NUMBER: 22477
; REFERENCE/DOCKET NUMBER: 29105/34166
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-906-791-2

Query Match 34.3%; Score 48; DB 3; Length 501;
Best Local Similarity 36.4%; Pred. No. 13;
Matches 12; Conservative 3; Mismatches 12; Indels 6; Gaps 1;

QY 2 GLRLGRKIAHCVKYGT-----VLRIIRA 28
Db 50 GLSLRHELVHGARYGPIMSGSGFTLRTVYA 82

RESULT 10
US-09-111-730-1
Sequence 1, Application US/09111730
Patent No. 6274359
GENERAL INFORMATION:
APPLICANT: Hideharu Anazawa
APPLICANT: Seiichi Shimada
APPLICANT: Seiichi Sugimoto
APPLICANT: Toshimasa Shinki
APPLICANT: Tatsuo Suda
APPLICANT: Yuzuru Ishimura
APPLICANT: Matsuhiko Hayashi
APPLICANT: Toshiaki Monkawa
APPLICANT: Tadashi Yoshida
APPLICANT: Shu Wakino
APPLICANT: Hiromichi Suzuki
TITLE OF INVENTION: 25-HYDROXYVITAMIN D3-1a-HYDROXYLASE AND DNA ENCODING THE HYDROXYLASE
FILE REFERENCE: 1074
CURRENT APPLICATION NUMBER: US/09/111,730
CURRENT FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 501
TYPE: PRT
ORGANISM: Rat
US-09-111-730-1

Query Match 34.3%; Score 48; DB 4; Length 501;
Best Local Similarity 36.4%; Pred. No. 13;
Matches 12; Conservative 3; Mismatches 12; Indels 6; Gaps 1;

QY 2 GLRLGRKIAHCVKYGT-----VLRIIRA 28
Db 50 GLSLRHELVHGARYGPIMSGSGFTLRTVYA 82

RESULT 11
US-08-782-997A-20
Sequence 20, Application US/08782997A
Patent No. 6030602
GENERAL INFORMATION:
APPLICANT: Legendre, Jean-Yves
APPLICANT: Supersaxo, Andreas
APPLICANT: Trzeciak, Arnold
TITLE OF INVENTION: Peptide Conjugates for Transfecting
TITLE OF INVENTION: Cells
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street

CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,997A
FILING DATE: 14-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 96100603.8
FILING DATE: 17-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kreisler, Lewis J
REGISTRATION NUMBER: 38,522
REFERENCE/DOCKET NUMBER: RAN 4600/73
TELECOMMUNICATION INFORMATION:
TELEPHONE: (973) 235-4387
TELEFAX: (973) 235-2363
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
US-08-782-997A-20

Query Match 33.6%; Score 47; DB 3; Length 34;
Best Local Similarity 36.8%; Pred. No. 0.87;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 3 LRLGRKIAHCVKYGT 21
Db 7 IEKMGRIIRNGIVKAGPAI 25

RESULT 12
US-08-039-557-3
Sequence 3, Application US/08039557
Patent No. 5714467
GENERAL INFORMATION:
APPLICANT: Boman, Hans G.
APPLICANT: Merrifield, Robert B.
APPLICANT: Andrew, David
TITLE OF INVENTION: ANTIBACTERIAL AND ANTIMALARIAL HYBRID
TITLE OF INVENTION: PEPTIDES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/039,557
FILING DATE: 26-MAR-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,462
FILING DATE: 05-FEB-1992

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/449,593
; FILING DATE: 12-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/336,777
; FILING DATE: 12-APR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-487-1684
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: YES
; FRAGMENT TYPE: N-terminal
; US-08-039-557-3

Query Match          33.6%; Score 47; DB 1; Length 35;
Best Local Similarity 36.8%; Pred. No. 0.9;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      3 LRLGRKIAHGKYGPTV 21
      : : | | | : | : | : | :
Db      8 IEMGRNIRNGIVKAGPAI 26

RESULT 13
US-08-782-997A-19
; Sequence 19, Application US/08782997A
; Patent No. 6030602
; GENERAL INFORMATION:
; APPLICANT: Legendre, Jean-Yves
; APPLICANT: Supersaxo, Andreas
; APPLICANT: Trzeciak, Arnold
; TITLE OF INVENTION: Peptide Conjugates for Transfecting
; TITLE OF INVENTION: Cells
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/782,997A
; FILING DATE: 14-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 96100603.8
; FILING DATE: 17-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38,522
; REFERENCE/DOCKET NUMBER: RAN 4600/73
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
```

```

; LENGTH: 35 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; US-08-782-997A-19

Query Match          33.6%; Score 47; DB 3; Length 35;
Best Local Similarity 36.8%; Pred. No. 0.9;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      3 LRLGRKIAHGKYGPTV 21
      : : | | | : | : | : | :
Db      8 IEMGRNIRNGIVKAGPAI 26

RESULT 14
US-08-782-997A-25
; Sequence 25, Application US/08782997A
; Patent No. 6030602
; GENERAL INFORMATION:
; APPLICANT: Legendre, Jean-Yves
; APPLICANT: Supersaxo, Andreas
; APPLICANT: Trzeciak, Arnold
; TITLE OF INVENTION: Peptide Conjugates for Transfecting
; TITLE OF INVENTION: Cells
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/782,997A
; FILING DATE: 14-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 96100603.8
; FILING DATE: 17-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38,522
; REFERENCE/DOCKET NUMBER: RAN 4600/73
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; US-08-782-997A-25

Query Match          33.6%; Score 47; DB 3; Length 35;
Best Local Similarity 36.8%; Pred. No. 0.9;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      3 LRLGRKIAHGKYGPTV 21
      : : | | | : | : | : | :
Db      8 IEMGRNIRNGIVKAGPAI 26
```

```

RESULT 15
US-08-749-066A-2
; Sequence 2, Application US/08749066A
; Patent No. 6303568
; GENERAL INFORMATION:
; APPLICANT: Jaynes, Jesse
; TITLE OF INVENTION: THERAPEUTIC ANTIMICROBIAL POLYPEPTIDES, THEIR USE AND METHODS OF
; FILE REFERENCE: 068370.0103
; PREPARATION
; CURRENT APPLICATION-NUMBER: US/08/749,066A
; CURRENT FILING DATE: 1996-11-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Hyalophora cecropia
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (35)..(35)
; OTHER INFORMATION: AMIDATION
US-08-749-066A-2

```

```

Query Match 33.6%; Score 47; DB 4; Length 35;
Best Local Similarity 36.8%; Pred. No. 0.9;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;
QY 3 LRLGRKTAHGVKKKPTV 21
: : | | | : | | :
Db 8 IEKMGRIIRNGIVKAGPAI 26

```

Search completed: July 12, 2002, 07:55:10
Job time: 117 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:56:58 ; Search time 95.47 Seconds
(without alignments)
28.182 Million cell updates/sec

Title: US-09-642-744B-27
Perfect score: 140
Sequence: 1 RGLRRGRKRIAGVKRKYGPVLRIRIA 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:
1: p1r1:
2: p1r2:
3: p1r3:
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	140	100.0	152	2 S68411	cathelin-related p
2	140	100.0	160	2 S68228	myeloid antimicrob
3	140	100.0	160	2 S68442..	cathelin-related p
4	35	39.3	273	2 F70158	conserved hypotet
5	54	38.6	367	1 K1EC6G	glutamate 5-Kinase
6	54	38.6	367	2 E90662	gamma-glutamate ki
7	54	38.6	367	2 C85513	glutamate 5-Kinase
8	54	38.6	367	1 AG0543	glutamate 5-Kinase
9	53.5	38.2	367	1 K1SEEM	glutamate 5-Kinase
10	53.5	38.2	367	2 AE0391	glutamate 5-Kinase
11	53	37.9	399	2 D81198	acetate kinase NMB
12	53	37.9	597	2 C69283	hypothetical prote
13	52	37.1	1524	2 C84721	hypothetical prote
14	51	36.4	399	2 F81775	acetyl kinase (EC
15	51	36.4	350	2 F90407	molybdopter in bios
16	50.5	36.1	292	2 B72659	deoxyhypusine synt
17	50	35.7	666	2 B73340	hypothetical prote
18	49.5	35.4	425	2 A12518	transposase alr732
19	49	35.0	139	2 T04310	strong gravity str
20	49	35.0	213	2 A30302	sodium channel pro
21	49	35.0	428	2 S35215	sodium channel pro
22	49	35.0	1689	2 S72467	sodium channel pro
23	49	35.0	1750	2 G84649	hypothetical prote
24	49	35.0	1820	2 A33299	sodium channel pro
25	49	35.0	2108	2 S72458	sodium channel pro
26	48	34.3	440	2 D64186	conserved hypotet
27	48	34.3	556	2 F97293	formate--tetrahydr
28	48	34.3	633	2 D75112	gla-tRNA amidotran
29	47.5	33.9	292	2 S73692	probable GTPase/GT

30	47	33.6	62	1 CKWKB	cectropin B precurs
31	47	33.6	450	2 A12324	hypothetical prote
32	47	33.6	520	2 B97464	hypothetical 54.3k
33	47	33.6	520	2 AD2682	conserved hypotet
34	47	33.6	801	1 T52538	1-phosphatidylinos
35	47	33.6	1053	2 D83206	probable RND efflu
36	46	32.9	35	1 CKA0BP	cectropin B - Chine
37	46	32.9	235	1 S76567	ribose-5-phosphate
38	46	32.9	423	2 S29184	sodium channel pro
39	46	32.9	492	2 D75619	cobrytic acid synth
40	46	32.9	1808	2 T47792	hypothetical prote
41	46	32.9	1951	2 S00320	sodium channel pro
42	46	32.9	1983	2 A60054	sodium channel pro
43	46	32.9	2005	2 B25019	sodium channel pro
44	46	32.9	2005	2 A46269	sodium channel alp
45	46	32.9	2009	2 A25019	sodium channel pro

ALIGNMENTS

```

RESULT 1
S68411
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68411
R.Mahoney, M.M., Lee, A.Y., Brezinski-Calliguri, D.J., Hutterer, K.M.
FEBS Lett. 377, 519-522, 1995
A>Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobia
A:Reference number: S68411; MUID:96140581
A:Accession: S68411
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-152 <MAH>
A:Cross-References: EMBL:X92757
C:Genetics:
A:Gene: SC5-2
C:Superfamily: cathelin: cystatin homology
F:1-21/Domain: signal sequence #status predicted <Sig>
F:14-122/Domain: cystatin homology <Cys>
F:21-123/Domain: propeptide #status predicted <PRO>
F:124-152/Product: cathelin-related protein 2 #status predicted <MAT>

Query Match      100.0%; Score 140; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 1e-12;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RGLRRGRKRIAGVKRKYGPVLRIRIA 28
Db      124 RGLRRGRKRIAGVKRKYGPVLRIRIA 151

RESULT 2
S68228
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68228
R.Bagella, L., Scofield, M., Zanetti, M.
FEBS Lett. 376, 225-228, 1995
A>Title: cDNA sequences of three sheep myeloid cathelicidins.
A:Reference number: S68228; MUID:96105386
A:Accession: S68228
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-160 <BAG>
A:Cross-References: EMBL:L46854; MUID:g1161248; PIDN:AAA85470.1; PID:g1161249
C:Superfamily: cathelin: cystatin homology
F:1-29/Domain: signal sequence #status predicted <Sig>
F:22-130/Domain: cystatin homology <Cys>

```

F:29-131/Domain: propeptide #status predicted <PRO>
F:132-160/Product: myeloid antimicrobial peptide 29 #status predicted <MAT>

Query Match 100.0%; Score 140; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGLRLGRKIAHGKKYKGPVLRIRIA 28
|||||
DB 132 RGLRLGRKIAHGKKYKGPVLRIRIA 159

RESULT 3

S68412
cathelin-related protein 1 precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68412
R:Mahoney, M.M.; Lee, A.Y.; Brezinski-Caliguri, D.J.; Hutner, K.M.
FEBS Lett. 377, 519-522, 1995
A:Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobial
A:Reference number: S68411; MUID:96140581
A:Accession: S68412
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-160 <MAH>
A:Cross-references: EMBL:X92758
C:Genetics:
A:Gene: SC5-1
C:Superfamily: cathelin; cystatin homology
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-130/Domain: cystatin homology <CTS>
F:30-111/Domain: propeptide #status predicted <PRO>
F:132-160/Product: cathelin-related protein 1 #status predicted <MAT>

Query Match 100.0%; Score 140; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.1e-12;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGLRLGRKIAHGKKYKGPVLRIRIA 28
|||||
DB 132 RGLRLGRKIAHGKKYKGPVLRIRIA 159

RESULT 4
F70158
conserved hypothetical protein BB0471 - Lyme disease spirochete

C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: F70158
R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
N.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 560-566, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943
A:Accession: F70158
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-273 <KLE>
A:Cross-references: GB:AE001151; GB:AE000783; NID:92688371; PIDN:AAC66832.1; PID:9268838
A:Experimental source: strain B31
C:Superfamily: hypothetical protein ymdb

Query Match 39.3%; Score 55; DB 2; Length 273;
Best Local Similarity 56.5%; Pred. No. 2.7;
Matches 13; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 GLRRLGRKIAHGKKYKGPVLRIRIA 24

DB 45 GLRGLGRKIAHFLKKYGDIVLRL 67
|||||

RESULT 5

KIECG
glutamate 5-kinase (EC 2.7.2.11) - Escherichia coli
N:Alternate names: gamma-glutamyl kinase
C:Species: Escherichia coli
C:Date: 31-Dec-1968 #sequence_revision 12-Dec-1997 #text_change 21-Jul-2000
C:Accession: C64749; A31001; C31001; B31001
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: C64749
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-367 <BLAT>
A:Cross-references: GB:AE000132; GB:U00096; NID:92367098; PIDN:AAC73346.1; PID:917864
A:Experimental source: Strain K-12, substrain MG1655
R:Deutch, A.H.; Rushlow, K.E.; Smith, C.J.
Nucleic Acids Res. 12, 6337-6355, 1984
A:Title: Analysis of the Escherichia coli proBA locus by DNA and protein sequencing.
A:Reference number: A93531; MUID:84297232
A:Accession: A31001
A:Molecule type: DNA
A:Residues: 1-142, 'A', 144-367
A:Cross-references: GB:D83536; NID:94902908; PIDN:BAA77911.1; PID:94902977
A:Accession: C31001
A:Molecule type: protein
A:Residues: 1-678-12 <DEP>
C:Comment: This enzyme catalyzes the ATP-dependent phosphorylation of glutamate to fo
C:Genetics:
A:Gene: probA
A:Map position: 6 min
C:Function:
A:Pathway: amino acid metabolism
A:Superfamily: glutamate 5-kinase
C:Keywords: phosphotransferase; proline biosynthesis
F:1-367/Product: glutamate 5-kinase #status experimental <MAT>

Query Match 38.6%; Score 54; DB 1; Length 367;
Best Local Similarity 59.1%; Pred. No. 4.9;
Matches 13; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

QY 7 GRKIAHGKKYKGPVLRIRIA 28
|||||
DB 320 GRDIAHGVSRIYSDALR--RIA 339

RESULT 6

E90662
gamma-glutamate kinase [imported] - Escherichia coli (strain O157:H7, substrain RIMD
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: E90662
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasakura, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: E90662
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-367 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033692.1; PID:913359726; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EGS0269
C:Superfamily: glutamate 5-kinase

```

Query Match      38.6%; Score 54; DB 2; Length 367;
Best Local Similarity 59.1%; Pred. No. 4.9;
Matches 13; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

OY 7 GRIAHGVKKYGPVLRIRIA 28
    || ||||| : | |||
Db 320 GRDIAHGVSRYNSDALR--RIA 339

RESULT 7
C85513
gamma-glutamate kinase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: C85513
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
  Iller, L.; Grobleck, E.J.; Davis, N.W.; Lim, A.; Dimadanta, E.; Potamoukis, K.; Apodaca,
  Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85513
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-367 <STO>
A:Cross-references: GB:AE005174; NID:q12513016; PIDN:AG54567.1; GSPDB:GN00145; UMGF:203
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: prob
C:Superfamily: glutamate 5-kinase

Query Match      38.6%; Score 54; DB 2; Length 367;
Best Local Similarity 59.1%; Pred. No. 4.9;
Matches 13; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

OY 7 GRIAHGVKKYGPVLRIRIA 28
    || ||||| : | |||
Db 320 GRDIAHGVSRYNSDALR--RIA 339

RESULT 8
AG0543
glutamate 5-kinase [imported] - Salmonella enterica subsp. enterica serovar Typhi (strai
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: This species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG0543
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
  th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
  , S.; Moule, S.; O'Gaora, P.
  Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
  A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
  A:Reference number: AB0502; PMID:11677608
A:Accession: AG0543
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-367 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD08791.1; PID:q16501607; GSPDB:GN00176
C:Genetics:
A:Gene: SRY0366
C:Superfamily: glutamate 5-kinase

Query Match      38.6%; Score 54; DB 2; Length 367;
Best Local Similarity 59.1%; Pred. No. 4.9;
Matches 13; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

OY 7 GRIAHGVKKYGPVLRIRIA 28
    || ||||| : | |||
Db 320 GRDIAHGVSRYNSDALR--RIA 339

```

```

RESULT 9
KISEEM
glutamate 5-kinase (EC 2.7.2.11) - Serratia marcescens
C:Species: Serratia marcescens
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 11-Jun-1999
C:Accession: A49753; S1644
R:Omori, K.; Suzuki, S.I.; Imai, Y.; Komatsubara, S.
  J. Gen. Microbiol. 137, 509-517, 1991
A:Title: Analysis of the Serratia marcescens proA operon and feedback control of pro
  A:Reference number: A49753; MUID:91237315
A:Accession: A49753
A:Molecule type: DNA
A:Residues: 1-367 <OMO>
A:Cross-references: GB:X53086; NID:g47251; PIDN:CAA37254.1; PID:g47253
C:Genetics:
A:Gene: prob
C:Superfamily: glutamate 5-kinase
C:Keywords: phosphotransferase; proline biosynthesis

Query Match      38.2%; Score 53.5; DB 1; Length 367;
Best Local Similarity 45.8%; Pred. No. 5.7;
Matches 11; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

OY 3 LRRL-GRIAHGVKKYGPVLRIRI 25
    : | | | | | : | | |
Db 315 LRNLGRDLAHGVSRYNSDAMRMI 338

RESULT 10
AE0391
glutamate 5-kinase (EC 2.7.2.11) [imported] - Yersinia pestis (strain C092)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE0391
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tiltball, R.W.; Holden, M.T.G.; Prentice, M
  deno-Farrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
  ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
  Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AE0391
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-367 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92457.1; PID:q15981158; GSPDB:GN00175
C:Genetics:
A:Gene: prob
C:Superfamily: glutamate 5-kinase
C:Keywords: phosphotransferase

Query Match      38.2%; Score 53.5; DB 2; Length 367;
Best Local Similarity 45.8%; Pred. No. 5.7;
Matches 11; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

OY 3 LRRL-GRIAHGVKKYGPVLRIRI 25
    : | | | | | : | | |
Db 315 LRNLGRDLAHGVSRYNSDAMRMI 338

RESULT 11
DB1198
acetate kinase NME0435 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: DB1198
R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
  Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
  ri, H.; Ojo, H.; Yamathavan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.
  Science 287, 1809-1815, 2000

```

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V&
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: D81198
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <REF>
A:Cross-references: GB:AE002400; GB:AE002098; NID:g7225659; PIDN:AAFA0873.1; PID:g722566
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0435
C:Superfamily: acetate kinase

Query Match 37.9%; Score 53; DB 2; Length 399;
Best Local Similarity 45.0%; Pred. No. 7.3;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRRIGRKIAHGKRYGPTVL 22
DB 83 IKAIGHRIAHGKRYSESVL 102

RESULT 12
C69283
hypothetical protein AF0267 - *Archaeoglobus fulgidus*
C:Species: *Archaeoglobus fulgidus*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Nov-1999
C:Accession: C69283
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uitterlinden, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343
A:Accession: C69283
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-597 <KLE>
A:Cross-references: GB:AE001087; GB:AE000782; NID:g2689410; PIDN:AAB90971.1; PID:g265037
C:Superfamily: *Archaeoglobus fulgidus* hypothetical protein AF0267

Query Match 37.9%; Score 53; DB 2; Length 597;
Best Local Similarity 45.5%; Pred. No. 11;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 GLRLGRKIAHGKRYGPTVL 23
DB 89 GLRLGRKIAHGKRYVSEFR 110

RESULT 13
G84721
hypothetical protein Atg31520 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G84721
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference number: A84420; MUID:20083487
A:Accession: G84721
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1524 <STO>
A:Cross-references: GB:AE002093; NID:g4582447; PIDN:AAD24831.1; GSPDB:GN00139
C:Genetics:
A:Gene: Atg31520

A:Map position: 2

Query Match 37.1%; Score 52; DB 2; Length 1524;
Best Local Similarity 55.6%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRRIGRKIAHGKRYGPT 20
DB 1483 IRRKGNKLAHVLAKYGCT 1500

RESULT 14
F81775
acetate kinase (EC 2.7.2.1) NMA2050 [imported] - *Neisseria meningitidis* (strain Z2491)
C:Species: *Neisseria meningitidis*
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: F81775
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mc
; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491
A:Reference number: A81775; MUID:20222556
A:Accession: F81775
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-399 <PAR>
A:Cross-references: GB:AL162758; GB:AL157959; NID:g7380672; PIDN:CAB85268.1; PID:g738
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: ackA2; NMA2050
C:Superfamily: acetate kinase
C:Keywords: Phosphotransferase

Query Match 36.4%; Score 51; DB 2; Length 399;
Best Local Similarity 45.0%; Pred. No. 14;
Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRRIGRKIAHGKRYGPTVL 22
DB 83 IOAVGHRIAHGKRYSESVL 102

RESULT 15
F90407
molybdopterin biosynthesis protein (moa-2) [imported] - *Sulfolobus solfataricus*
C:Species: *Sulfolobus solfataricus*
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: F90407
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Cl
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A89139
A:Accession: F90407
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-550 <KUR>
A:Cross-references: GB:AE006641; NID:g13815669; PIDN:AAK42517.1; GSPDB:GN00155
C:Genetics:
A:Gene: moa-2

Query Match 36.4%; Score 51; DB 2; Length 550;
Best Local Similarity 44.4%; Pred. No. 19;
Matches 12; Conservative 6; Mismatches 7; Indels 2; Gaps 2;

QY 1 RGRLGRKIAHGKRYGPTVLRII 25
DB 268 RAIKELGSIIHGIIKIKIPGKPTILGV 294

Fri Jul 12 08:56:21 2002

us-09-642-744b-27.rpr

Page 5

Search completed: July 12, 2002, 07:56:59
Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:04:43 ; Search time 46.15 Seconds
(without alignments)
23.492 Million cell updates/sec

Title: US-09-642-744B-27
Perfect score: 140
Sequence: 1 RGLRLGRKRIANGVKRKYPTVLRIRIA 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	160	SC51_SHEEP	P49928 ovis aries
2	140	100.0	160	SC52_SHEEP	P49929 ovis aries
3	82	58.6	159	MB28_BOVIN	P54229 bos taurus
4	54	38.6	367	PROB_ECOLI	P07005 escherichia
5	53.5	38.2	367	PROB_SERNA	P17856 serratia ma
6	53	37.9	597	Y267_ARCFU	O29722 archaeoglob
7	50.5	36.1	292	DHYS_AERPE	O9y478 aeropyrum p
8	49	35.0	139	GRAL_ORYSA	O07077 oryza sativ
9	49	35.0	2131	CINA_DROME	P35500 drosophila
10	48	34.3	440	YDJN_HAEIN	P45079 haemophilus
11	48	34.3	501	CP2B_RAT	O35132 rattus norv
12	48	34.3	507	CP2B_MOUSE	O35084 mus musculu
13	47.5	33.9	449	EMGA_MYCPN	P75309 mycoplasma
14	47	33.6	62	CECB_HYACE	P01508 hyalophora
15	47	33.6	326	LDH_TOXGO	O27797 toxoplasma
16	47	33.6	801	VP34_SCHPO	P50520 schizosacch
17	47	33.6	1951	CIN3_HUMAN	O9y476 homo sapien
18	46	32.9	33	CECB_ANTPE	P01509 antheraea p
19	46	32.9	235	REPA_SYNY3	O51766 synecocyst
20	46	32.9	1951	CIN3_RAT	P08104 rattus norv
21	46	32.9	2005	CIN2_HUMAN	O99250 homo sapien
22	46	32.9	2005	CIN2_RAT	P04775 rattus norv
23	46	32.9	2009	CIN1_HUMAN	P35498 homo sapien
24	46	32.9	2009	CIN1_RAT	P04774 rattus norv
25	45.5	32.5	524	GUAA_YEAST	P38625 saccharomyc
26	45	32.1	63	CECB_BOOMO	P04142 bombyx mori
27	45	32.1	112	MYRL_MYRPI	O07532 myrmecia pi
28	45	32.1	370	YXER_BACST	P54957 bacillus su
29	45	32.1	432	CHAA_COMTE	O44256 camomonas t
30	45	32.1	537	SYE_TREPA	O81679 treponema p
31	45	32.1	572	PTL_LISTIN	O92419 listeria in
32	45	32.1	572	PTL_LISMO	O31149 listeria mo
33	45	32.1	578	PT1_BACSP	O63018 bacillus sp

34	45	32.1	578	1	PT1_BACST	P42014 bacillus st
35	45	32.1	598	1	ILV3_SCHPO	O10318 schizosacch
36	45	32.1	635	1	YE64_PYRHO	O59133 pyrococcus
37	45	32.1	730	1	ROM_HUMAN	P52272 homo sapien
38	44.5	31.8	216	1	YE42_SCHPO	O13962 schizosacch
39	44.5	31.8	418	1	NAGA_CAEEL	P34480 caenorhabdi
40	44.5	31.8	702	1	HELS_AERPE	O9y478 aeropyrum p
41	44	31.4	63	1	CE3F_HYPCU	P50722 hyphantria
42	44	31.4	231	1	Y333_MYCPN	P75322 mycoplasma
43	44	31.4	256	1	ATPE_HUMAN	P24539 homo sapien
44	44	31.4	369	1	V736_HELPJ	O92147 helicobacte
45	44	31.4	369	1	V736_HELPJ	O25436 helicobacte

ALIGNMENTS

RESULT ID	SC51_SHEEP	STANDARD	PRT	160 AA.
AC	P49928:			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Cathelin-related peptide SC5 precursor 1 (Antibacterial peptide SMAP-29) (Myeloid antibacterial peptide SMAP-29).			
DE	Ovis aries (Sheep): Chordata; Craniata; Vertebrata; Euteleostomi; OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; OC Bovidae; Caprinae; Ovis.			
OX	NCBI_TaxID=9940;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Bone marrow;			
RX	MEDLINE=96140581; PubMed=8549789;			
RA	Mahoney M.M., Lee A.Y., Brezinski-Caliguri D.J., Huttner K.M.;			
RT	"Molecular analysis of the sheep cathelin family reveals a novel antimicrobial peptide."			
RT	FEBS Lett. 377:519-524(1995)			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RA	Huttner K.M., Mahoney M.M.;			
RL	Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.			
CC	- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.			
CC	- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.			
CC	- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	EMBL: X92757; CA63412.1; -			
DR	EMBL: U06000; AAB49715.1; -			
DR	InterPro: IPR001894; Cathelicidin.			
DR	Pfam: PF00666; Cathelicidins; 1.			
DR	ProDom: PD001838; Cathelicidins; 1.			
DR	PROSITE: PS00946; CATHELICIDINS; 1.			
DR	PROSITE: PS00947; CATHELICIDINS_2; 1.			
FT	Antibiotic; Signal.			
FT	SIGNAL	1	29	POTENTIAL.
FT	PROPEP	30	131	BY SIMILARITY.
FT	CHAIN	132	160	CATHELIN-RELATED PEPTIDE SC5.
FT	MOD_RES	30	30	PYRROLIDONE CARBOXYLIC ACID
FT				(BY SIMILARITY).
FT	DISULFD	86	97	BY SIMILARITY.
FT	DISULFD	108	125	BY SIMILARITY.
FT	SEQUENCE	160 AA;	17786 MW;	BD9B3859CA32C249 CRC64;

Query Match 100.0%; Score 140; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 5.9e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGLRLGRKRIAGVKKYGPVLRIRIA 28
Db 132 RGLRLGRKRIAGVKKYGPVLRIRIA 159

RESULT 2

SC52_SHEEP STANDARD; PRT: 160 AA.
AC P49929;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cathelin-related peptide SC5 precursor 2 (Antibacterial peptide SMAP-29) (Myeloid antibacterial peptide SMAP-29).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=96140581; PubMed=8549789;
RA Maoney M.M., Lee A.Y., Brezinski-Caliguri D.J., Hutter K.M.;
RT "Molecular analysis of the sheep cathelin family reveals a novel antimicrobial peptide.";
RL FEBS Lett. 377:519-522(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=96105386; PubMed=7498547;
RA Bagella L., Scocchi M., Zanetti M.;
RT "cDNA sequences of three sheep myeloid cathelicidins.";
RL FEBS Lett. 376:225-228(1995).
CC -1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X92758; CAA63413.1; -;
DR InterPro: IPR001894; Cathelicidin.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00666; Cathelicidins; 1.
DR Prodom: PD001838; Cathelicidins; 1.
DR SMART: SM00043; CY; 1.
DR PROSITE: PS00946; CATHELICIDINS_1; 1.
DR PROSITE: PS00947; CATHELICIDINS_2; 1.
KW Antibiotic; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 131 BY SIMILARITY.
FT CHAIN 132 160 CATHELIN-RELATED PEPTIDE SC5.
FT MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID
(BY SIMILARITY).
FT DISULFID 86 97 BY SIMILARITY.
FT DISULFID 108 125 BY SIMILARITY.
FT CONFLICT 28 28 S -> R (IN REF. 2).
SQ SEQUENCE 160 AA; 17742 MW; 4FB98A09355B51F CRC64;

QY 1 RGLRLGRKRIAGVKKYGPVLRIRIA 28
Db 132 RGLRLGRKRIAGVKKYGPVLRIRIA 159

RESULT 3

MB28_BOVIN STANDARD; PRT: 159 AA.
AC P54229;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Antibacterial peptide BMAP-28 precursor (Myeloid antibacterial peptide BMAP-28).
GN BMAP28.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=97067059; PubMed=8910461;
RA Skerlavaj B., Gennaro R., Bagella L., Merluzzi L., Risso A.,
Zanetti M.;
RT "Biological characterization of two novel cathelicidin-derived peptides and identification of structural requirements for their antimicrobial and cell lytic activities.";
RL J. Biol. Chem. 271:28375-28381(1996).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=96300243; PubMed=8706679;
RA Storici P., Tossi A., Lenarcic B., Romeo D.;
RT "Purification and structural characterization of bovine cathelicidins, precursors of antimicrobial peptides.";
RL Eur. J. Biochem. 238:769-776(1996).
CC -1- FUNCTION: EXERTS A POTENT ANTIMICROBIAL ACTIVITY AGAINST GRAM-NEGATIVE AND GRAM-POSITIVE BACTERIA, INCLUDING METHICILLIN-RESISTANT STAPHYLOCOCCUS AUREUS, AND FUNGI.
CC -1- MASS SPECTROMETRY: MW=14414; MW_ERR=1; METHOD=Electrospray;
CC RANGE=30-159.
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X97609; CAA66208.1; -;
DR InterPro: IPR001894; Cathelicidin.
DR Pfam: PF00666; Cathelicidins; 1.
DR Prodom: PD001838; Cathelicidins; 1.
DR PROSITE: PS00946; CATHELICIDINS_1; 1.
DR PROSITE: PS00947; CATHELICIDINS_2; 1.
KW Antibiotic; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 131 BY SIMILARITY.
FT CHAIN 132 159 ANTIANTIBACTERIAL PEPTIDE BMAP-28.
FT MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 86 97 BY SIMILARITY.
FT DISULFID 108 125 BY SIMILARITY.
SQ SEQUENCE 159 AA; 17616 MW; 89B7CB46C5EEC367 CRC64;

Query Match 100.0%; Score 140; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 5.9e-14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 58.6%; Score 82; DB 1; Length 159;
Best Local Similarity 65.4%; Pred. No. 2.9e-05;
Matches 17; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

OY      2  GRLRLGRKIAHGKYGKPYTLRIIRI 27
        ||| |||| |||| ||||
DB      133 GRLSLGRKILRLAKKKYPIVPIIRI 158

RESULT  4
PROB_ECOLI  STANDARD;  PRT;  367 AA.
AC      P07005; P78293;
DT      01-APR-1988 (Rel. 07, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      Glutamate 5-kinase (EC 2.7.2.11) (Gamma-glutamyl kinase) (GK).
GN      PROB OR B0242 OR Z0303 OR ECS0269.
OS      Escherichia coli, and
OS      Escherichia coli O157:H7.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562, 83334;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 1-6 AND 8-12.
RX      MEDLINE=84297232; PubMed=6089111;
RA      Deutch A.H., Rushlow K.E., Smith C.J.;
RT      "Analysis of the Escherichia coli proBA locus by DNA and protein
RT      sequencing."
RL      Nucleic Acids Res. 12:6337-6355(1984).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      STRAIN-K12 / MG1655;
RA      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12."
RL      Science 277:1453-1474(1997).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      STRAIN-K12 / WJ110;
RA      Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA      Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA      Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;
RT      "Systematic sequencing of the Escherichia coli genome: analysis of the
RT      4.0 - 6.0 min (189,987 - 281,416bp) region."
RL      Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A.
RA      Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA      Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA      Lashkari D., Lew H., Lin D., Nemeth A., Oefner P., Roberts D.,
RA      Davis R.W.;
RT      Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      SEQUENCE FROM N.A.
RX      STRAIN-O157:H7 / EDL933 / ATCC 700927;
RA      MEDLINE=21074935; PubMed=11206551;
RA      Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA      Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA      Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA      Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
RA      Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA      Welch R.A., Blattner F.R.;
RT      "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL      Nature 409:529-533(2001).
RN      [6]
RP      SEQUENCE FROM N.A.
RX      STRAIN-O157:H7 / RIMD 0509952;
RA      MEDLINE=21156231; PubMed=11258796;
RA      Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA      Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA      Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA      Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT      "Complete genome sequence of enterohaemorrhagic Escherichia coli

```

```

RT      O157:H7 and genomic comparison with a laboratory strain K-12."
RL      DNA Res. 8:11-22(2001).
RN      [7]
RP      SEQUENCE OF 1-62 FROM N.A.
RX      STRAIN-K12;
RA      MEDLINE=83189086; PubMed=6341601;
RA      Overbeek N., Bergmans H., van Mansfeld F., Lugtenberg B.;
RT      "Complete nucleotide sequence of phoE, the structural gene for the
RT      phosphate limitation inducible outer membrane pore protein of
RT      Escherichia coli K12."
RL      J. Mol. Biol. 163:513-532(1983).
CC      -1- FUNCTION: Catalyzes the transfer of a phosphate group to glutamate
CC      to form glutamate 5-phosphate which rapidly cyclizes to 5-
CC      oxoproline.
CC      -1- CATALYTIC ACTIVITY: ATP + L-glutamate -> ADP + L-glutamate 5-
CC      phosphate.
CC      -1- PATHWAY: Proline biosynthesis pathway; first step.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE GLUTAMATE 5-KINASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X00786; CA25363.1; -
DR      EMBL; AE000132; AAC73346.1; -
DR      EMBL; D83536; BAA77911.1; -
DR      EMBL; U70214; BAA08662.1; -
DR      EMBL; AE005202; AAG54567.1; -
DR      EMBL; AP002551; BAB33692.1; -
DR      EMBL; V00316; CAA23604.1; -
DR      PIR; A31001; KIECEG.
DR      EcoGene; EGI0768; prob.
DR      InterPro; IPR001048; Aakkinase.
DR      InterPro; IPR001057; Glut_5_kinase.
DR      InterPro; IPR002478; PUA.
DR      Pfam; PF00656; aakkinase; 1.
DR      Pfam; PF01472; PUA; 1.
DR      PRINTS; PR00474; GLO5KINASE.
DR      SMART; SM00359; PUA; 1.
DR      PROSITE; PS00902; GLUTAMATE_5_KINASE; 1.
KW      Transferrase; Kinase; Proline biosynthesis; Complete proteome.
FT      DOMAIN 159 190
FT      ROLE IN THE FORMATION OF CARBOXYL
FT      PHOSPHATES OF AMINO ACIDS (POTENTIAL).
FT      CONFLICT 143 143 E -> A (IN REF. 1 AND 3).
FT      SEQUENCE 367 AA; 39056 MW; 5541040202EBCCEB CRC64;
SQ

Query Match 38.6%; Score 54; DB 1; Length 367;
Best Local Similarity 59.1%; Pred. No. 1;
Matches 13; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

OY      7  GRKIAHGKYGKPYTLRIIRIA 28
        ||| |||| |||| ||||
DB      320 GRDIAHGVSRIYNSDALR-RIA 339

RESULT  5
PROB_SERMA  STANDARD;  PRT;  367 AA.
AC      P17836;
DT      01-AUG-1990 (Rel. 15, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      Glutamate 5-kinase (EC 2.7.2.11) (Gamma-glutamyl kinase) (GK).
GN      PROB.
OS      Serratia marcescens.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Serratia.

```



```

ox NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SR41;
RA MEDLINE=91237315; PubMed=1851803;
RA Omori K., Suzuki S., Imai Y., Komatsubara S.;
RA "Analysis of the Serratia marcescens proBA operon and feedback
RT control of proline biosynthesis."
RL J. Gen. Microbiol. 137:509-517(1991).
RN [2]
RP MUTANT.
RX MEDLINE=92268849; PubMed=13169377;
RA Omori K., Suzuki S., Imai Y., Komatsubara S.;
RA "Analysis of the mutant proBA operon from a proline-producing strain
RT of Serratia marcescens."
RL J. Gen. Microbiol. 138:693-699(1992).
CC -I- FUNCTION: Catalyzes the transfer of a phosphate group to glutamate
CC to form glutamate-5-phosphate which rapidly cyclizes to 5-
CC oxoproline.
CC -I- CATALYTIC ACTIVITY: ATP + L-glutamate -> ADP + L-glutamate 5-
CC phosphate.
CC -I- ENZYME REGULATION: PROLINE-MEDIATED FEEDBACK INHIBITION.
CC -I- PATHWAY: Proline biosynthesis pathway; first step.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: BELONGS TO THE GLUTAMATE 5-KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X53086; CAA37254.1; -
DR EMBL: D90351; BAA14364.1; -
DR PIR: S11644; KISEEM.
DR InterPro: IPR001048; Aakkinase.
DR InterPro: IPR001057; Glut_5_kinase.
DR InterPro: IPR002478; PUA.
DR Pfam: PF00696; aakkinase; 1.
DR Pfam: PF01472; PUA; 1.
DR PRINTS: PR00474; GLU5KINASE.
DR SMART: SM00359; PUA; 1.
DR PROSITE: PS00902; GLUTAMATE_5_KINASE; 1.
KW Transferase; Kinase; Proline biosynthesis.
FT DOMAIN 159 190
FT PROLINE IN THE FORMATION OF CARBOXYL
FT VARIANT 117 117
FT A -> V (IN PROLINE HYPERPRODUCING MUTANT;
FT 700-X LESS SENSITIVE TO PRO INHIBITION).
FT FT
SQ SEQUENCE 367 AA; 39169 MW; 68A2AE8F78353AD5 CRC64;

Query Match 38.2%; Score 53.5; DB 1; Length 367;
Best Local Similarity 45.8%; Pred. No. 1.2;
Matches 11; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

OY 3 LRR-L-GRIANGVKKYGPVTRLI 25
DB 315 IRLNLTGRDLAHGVSRYNSDAMRMI 338

RESULT 6
Y267_ARCFU ID Y267_ARCFU STANDARD; PRT; 597 AA.
AC 029972;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF0267.
GN AF0267.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;

```

```

OX Archaeoglobus.
RX NCBI_TaxID=2234;
RN [1]
RC STRAIN=VC-16 / DSM 4304 / ATCC 49559;
RX MEDLINE=98049343; PubMed=9389475;
RA Klank H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Kirsch S.E., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Giodok A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utechtack T.,
RA Cotton M.D., Spilligs T., Arlanch P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC DR EMBL; AF001087; MAB90971.1; -.
DR TIGR; AF0267; -.
KM Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 4 23 POTENTIAL.
FT TRANSMEM 209 231
SQ SEQUENCE 597 AA; 66250 MW; E2BCA81C0E3B2540 CRC64;

Query Match 37.9%; Score 53; DB 1; Length 597;
Best Local Similarity 45.5%; Pred. No. 2.4;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 GLRLGRKRIANGVKKYKPTVLR 23
II:|:::|:|::|:|:|:|
Db 89 GLTKLGKRVAKNGKYPSEPR 110

RESULT 7
DHYS_AERPE STANDARD; PRT: 292 AA.
AC OYEY72;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Probable deoxyhypusine synthase (EC 2.5.1.46) (DHS).
DN DYS OR APE0698
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcates; Desulfurococcaceae;
OC Aeropyrum.
OC NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1.
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jli-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RT DNA Res. 6:83-101(1999).
RL -1- FUNCTION: CATALYZES THE NAD-DEPENDENT OXIDATIVE CLEAVAGE OF

```

CC SPERMIDINE AND THE SUBSEQUENT TRANSFER OF THE BUTYLAMINE MOIETY OF
CC SPERMIDINE TO THE EPSILON-AMINO GROUP OF A SPECIFIC LYSINE RESIDUE
CC OF THE EIF-5A PRECURSOR PROTEIN TO FORM THE INTERMEDIATE
CC DEOXYHYPOSINE RESIDUE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: [eif5a-precursor]-lysine + spermidine =
CC [eif5a-precursor]-deoxyhypusine + propane-1,3-diamine.
CC -1- COFACTOR: NAD (BY SIMILARITY).
CC -1- PATHWAY: Hypusine biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE DEOXYHYPOSINE SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000060; BAA79674.1; -
DR HSSP: P49366; 1DHS.
DR InterPro: IPR002773; DS.
DR Pfam: PF01916; DS: 1.
DR ProDom: PD007730; DS: 1.
KW Hypusine biosynthesis; Transferase; NAD; Complete proteome.
SQ SEQUENCE 292 AA; 31830 MW; 75E72B5DBD19D38E CRC64;

Query Match 36.1%; Score 50.5; DB 1; Length 292;
Best Local Similarity 46.2%; Pred. No. 2.7;
Matches 12; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 1 GRLRLGRKIAHGKRYGPTVLRIR 26
| : | | : | : | | | : |
Db 95 RGVHRLG-NVFIPESYGPLVERFR 119

RESULT 8
GRAL_ORYSA STANDARD; PRT: 139 AA.
AC 007077;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Gravity specific protein GSC 233.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihaldoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NIPPONBARE; TISSUE=Callus;
RX MEDLINE=93199820; Pubmed=1363521;
RA Kwon S., Kikuchi S., Oono K.;
RT "Molecular cloning and characterization of gravity specific cDNA in
RT rice (Oryza sativa L.) suspension callus.",
RT Jpn. J. Genet. 67:335-348(1992).
RL 1-1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN CALLUS.
CC -1- INDUCTION: BY GRAVITY STRESS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D11137; BAA01911.1; -
DR EMBL: S56877; AAB25853.1; -
SQ SEQUENCE 139 AA; 15563 MW; A1C302F1988AC864 CRC64;

Query Match 35.0%; Score 49; DB 1; Length 139;
Best Local Similarity 43.5%; Pred. No. 2.2;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 GRLRLGRKIAHGKRYGPTVLRIR 24
| : | | : | : | | | : |
Db 101 GPRLLGAEIVHGVEVQGRLLRL 123

RESULT 9
CINA_DROME STANDARD; PRT: 2131 AA.
ID CINA_DROME
AC P35500; O15994; Q24082; Q24083; Q24084; Q24528; Q24529; Q24530;
AC Q24531; Q24532; Q9VXF7; P92137;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Sodium channel protein para (paralytic protein).
GN PARA OR CG9907.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=CANTON-S;
RX MEDLINE=94238327; Pubmed=8182428;
RA Thackeray J.R., Ganetzky B.;
RT "Developmentally regulated alternative splicing generates a complex
RT array of Drosophila para sodium channel isoforms.",
RL J. Neurosci. 14:2569-2578(1994).
RN (2)
RP SEQUENCE FROM N.A. (ISOFORM D).
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Lin X., Mattei I., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).
RN [3] SEQUENCE OF 1-1862 FROM N.A., FUNCTION, AND ALTERNATIVE SPLICING.
RC STRAIN-CANTON-S; TISSUE-Head;
RX MEDLINE-89376565; PubMed-2550145;
RA Loughney K., Kreber R., Ganetzky B.;
RT "Molecular analysis of the para locus, a sodium channel gene in
Drosophila.";
RL Cell 58:1143-1154(1989).
RN [4] SEQUENCE OF 1683-1895 FROM N.A.
RX MEDLINE-89184571; PubMed-2538830;
RA Ramaswami M., Tanouye M.A.;
RT "Two sodium-channel genes in Drosophila: implications for channel
diversity.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:2079-2082(1989).
RN [5] SEQUENCE OF 51-61 FROM N.A.
RA Tanaka Y.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [6] SEQUENCE OF 51-61 FROM N.A.
RA Tanaka Y., Yagi Y., Gamo S.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [7] ALTERNATIVE SPLICING, AND DEVELOPMENTAL STAGE.
RX MEDLINE-96042905; PubMed-8536968;
RA Thackeray J.R., Ganetzky B.;
RT "Conserved alternative splicing patterns and splicing signals in the
Drosophila sodium channel gene para.";
RL Genetics 141:203-214(1995).
CC -1- FUNCTION: THIS PROTEIN MEDIATES THE VOLTAGE-DEPENDENT SODIUM ION
PERMEABILITY OF EXCITABLE MEMBRANES. ASSUMING OPENED OR CLOSED
CONFORMATION IN RESPONSE TO THE VOLTAGE DIFFERENCE ACROSS THE
MEMBRANE, THE PROTEIN FORMS A SODIUM-SELECTIVE CHANNEL THROUGH
WHICH NA++ IONS MAY PASS IN ACCORDANCE WITH THEIR ELECTROCHEMICAL
GRADIENT.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: At least 7 isoforms: A, B, C, D (shown
here), E, F24 and F30; are produced by alternative splicing.
CC Further isoforms have been identified but not yet sequenced. These
have different combinations of the optional exons A, B, C, D, E
and F. Isoforms always have either exon C or D as these encode
segment S4. Sequence identity to para from D.virilis suggests
there may also be optional exons H and I.
CC -1- DEVELOPMENTAL STAGE: Isoforms A and B are seen in embryos and
adults. Isoforms D, F24 and F30 are predominant in embryos and
isoforms C and E predominant in adults.
CC -1- DOMAIN: THE SEQUENCE CONTAINS 4 INTERNAL REPEATS, EACH WITH 5
HYDROPHOBIC SEGMENTS (S1,S2,S3,S5,S6) AND ONE POSITIVELY CHARGED
SEGMENT (S4). SEGMENTS S4 ARE PROBABLY THE VOLTAGE-SENSORS AND ARE
CHARACTERIZED BY A SERIES OF POSITIVELY CHARGED AMINO ACIDS AT
EVERY THIRD POSITION.
CC -1- SIMILARITY: TO OTHER SODIUM CHANNEL PROTEINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL, U26713; AAA98541.1; -
DR EMBL, U26714; AAA98542.1; -
DR EMBL, U26715; AAA98543.1; -
DR EMBL, U26716; AAA98544.1; -
DR EMBL, U26716; AAA98545.1; -
DR EMBL, U26716; AAA98546.1; -
DR EMBL, U26716; AAA98547.1; -
DR EMBL, U26716; AAA98548.1; -
DR EMBL, U26717; AAA98549.1; -
DR EMBL, AE003502; AAF48617.1; ALT_SEQ.

DR EMBL, M32078; AAB59190.1; -
DR EMBL, M32078; AAB59191.1; -
DR EMBL, M32078; AAB59192.1; -
DR EMBL, M32078; AAB59193.1; -
DR EMBL, M32078; AAB59194.1; -
DR EMBL, M32078; AAB59195.1; -
DR EMBL, AB008113; BAA22890.1; -
DR EMBL, AB035812; BAA68526.1; -
DR PIR, A33299; A33299.
DR HSSP; P04002; 1MFA.
DR FLYBase; FBgn0003036; para.
DR InterPro; IPR002111; Cat_channel_TrpL.
DR InterPro; IPR000636; Cation_chan_noh_lig.
DR InterPro; IPR001882; Channel_pore_Ca_Na.
DR Pfam; PF00520; Ion_trans_4.
DR PRINTS; PR00170; NACHANNEL.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Glycoprotein; Repeat; Alternative splicing; Phosphorylation.
FT REPEAT 134 467
FT REPEAT I.
FT REPEAT II.
FT REPEAT III.
FT REPEAT IV.
FT TRANSMEM 149 172 S1 OF REPEAT I.
FT TRANSMEM 181 199 S2 OF REPEAT I.
FT TRANSMEM 213 231 S3 OF REPEAT I.
FT TRANSMEM 238 257 S4 OF REPEAT I.
FT TRANSMEM 274 297 S5 OF REPEAT I.
FT TRANSMEM 406 427 S6 OF REPEAT I.
FT TRANSMEM 813 837 S1 OF REPEAT II.
FT TRANSMEM 849 873 S2 OF REPEAT II.
FT TRANSMEM 881 900 S3 OF REPEAT II.
FT TRANSMEM 907 926 S4 OF REPEAT II.
FT TRANSMEM 942 963 S5 OF REPEAT II.
FT TRANSMEM 1014 1041 S6 OF REPEAT II.
FT TRANSMEM 1297 1320 S1 OF REPEAT III.
FT TRANSMEM 1335 1359 S2 OF REPEAT III.
FT TRANSMEM 1366 1387 S3 OF REPEAT III.
FT TRANSMEM 1392 1413 S4 OF REPEAT III.
FT TRANSMEM 1433 1454 S5 OF REPEAT III.
FT TRANSMEM 1534 1560 S6 OF REPEAT III.
FT TRANSMEM 1615 1638 S1 OF REPEAT IV.
FT TRANSMEM 1650 1673 S2 OF REPEAT IV.
FT TRANSMEM 1680 1703 S3 OF REPEAT IV.
FT TRANSMEM 1714 1735 S4 OF REPEAT IV.
FT TRANSMEM 1751 1773 S5 OF REPEAT IV.
FT TRANSMEM 1836 1860 S6 OF REPEAT IV.
FT MOD_RES 553 553 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT MOD_RES 570 570 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT CARBOHYD 313 313 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 682 682 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 982 982 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1180 1180 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1463 1463 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1482 1482 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1862 1862 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 2047 2047 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPIC 555 575 MISSING (IN ISOFORM A).
FT VARSPIC 763 770 MISSING (IN ISOFORM B).
FT VARSPIC 914 967
FT FT

Query Match 35.0%; Score 49; DB 1; Length 2131;
Best local similarity 40.0%; Pred. No. 34;
Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1;
QY 6 LGRKIHGVKKY--GPTVLIIRIA 28
DB 1695 LGLVLSDIIEKYFVSPILLRVRA 1719

```

RESULT 10
ID YDUN_HAEIN STANDARD: PRT: 440 AA.
AC P45079;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical symporter H1154.
GN H1154.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedbloom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrman J.L., Geoghagen N.S.M.,
RA Grehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE SODIUM:DICARBOXYLATE SYMPORTER FAMILY
CC (SDF). STRONG, TO E.COLI YDUN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U3795; AAC22809.1; -
CC TIGR: H1154; -
DR InterPro: IPR001991; Na_dicarboxyl_symp.
DR Pfam: PF00375; SDF: 1.
DR PROSITE: PS00713; NA_DICARBOXYL_SYP_1; FALSE_NEG.
DR Hypothetical protein; Transmembrane; Transport; Complete proteome.
KM
FT TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 70 90 POTENTIAL.
FT TRANSMEM 101 121 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
FT TRANSMEM 226 246 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 343 363 POTENTIAL.
FT TRANSMEM 366 386 POTENTIAL.
FT TRANSMEM 389 409 POTENTIAL.
SQ SEQUENCE 440 AA; 46491 MW; F9298839559FD06A CRC64;

```

```

CP2B_RAT
ID CP2B_RAT STANDARD: PRT: 501 AA.
AC 035132; 035076;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 25-hydroxyvitamin D-1 alpha hydroxylase, mitochondrial precursor
DE (EC 1.14.-.-) (25-OHD-1 alpha-hydroxylase) (25-hydroxyvitamin D3 1-
DE alpha-hydroxylase) (VD3 1A hydroxylase) (P450C1 alpha) (P450VD1-
DE alpha).
GN CYP27B1 OR CYP27B.
OS Rattus norvegicus (rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=97472245; PubMed=9333115;
RA St Arnaut R., Messerlian S., Moir J.M., Omdahl J.L., Glorieux F.H.;
RT "The 25-hydroxyvitamin D 1-alpha-hydroxylase gene maps to the
RL pseudovitamin D-deficiency rickets (PDDR) disease locus."
RN J. Bone Miner. Res. 12:1552-1559(1997).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=98058745; PubMed=9371776;
RA Shinkai T., Shimada H., Wakino S., Anazawa H., Hayashi M., Saruta T.,
RA Deluca H.F., Suda T.;
RT "Cloning and expression of rat 25-hydroxyvitamin D3-1alpha-hydroxylase
RT cDNA."
RL Proc. Natl. Acad. Sci. U.S.A. 94:12920-12925(1997).
CC -1- FUNCTION: CATALYZES THE CONVERSION OF 25-HYDROXYVITAMIN D3
CC (25(OH)D) TO 1-ALPHA,25-DIHYDROXYVITAMIN D3 (1,25(OH)2D) PLAYS AN
CC IMPORTANT ROLE IN CRUCIAL ROLE IN NORMAL BONE GROWTH, CALCIUM
CC METABOLISM, AND TISSUE DIFFERENTIATION.
CC -1- PATHWAY: SECOND STEP IN THE CONVERSION OF VITAMIN D(3) INTO THE
CC ACTIVE FORM (1-ALPHA,25-DIHYDROXYVITAMIN D(3)).
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- TISSUE SPECIFICITY: KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AF000139; AAB86461.1; -
CC EMBL: AB001992; BAA3271.1; -
CC HSSP: P00189; 1SCC.
DR InterPro: IPR001128; CYL_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KM Electron transport; Oxidoreductase; Monooxygenase; Membrane;
KM Heme; Mitochondrion; Transl.
KW
FT CHAIN 1 501 MITOCHONDRION (POTENTIAL).
FT TRANS 25-HYDROXYVITAMIN D-1 ALPHA HYDROXYLASE.
FT BINDING 448 448 HEME (BY SIMILARITY).
FT CONFLICT 13 13 H -> D (IN REF. 1).
FT CONFLICT 55 55 H -> D (IN REF. 1).
FT CONFLICT 103 112 FSSWSEHRR -> SHLOSTIVAS (IN REF. 1).
FT CONFLICT 119 119 L -> W (IN REF. 1).
FT CONFLICT 129 144 RLRSLAPLLRPQAA -> EAPKSPGASPTSSS (IN
FT REF. 1).
FT CONFLICT 201 201 G -> R (IN REF. 1).
FT CONFLICT 251 251 D -> N (IN REF. 1).
FT CONFLICT 288 288 H -> D (IN REF. 1).
FT CONFLICT 305 305 T -> R (IN REF. 1).
FT CONFLICT 372 374 RLY -> MLD (IN REF. 1).

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:03:53 : Search time 174.7 Seconds
(without alignments)
27.727 Million cell updates/sec

Title: US-09-642-744B-27
Perfect score: 140
Sequence: 1 RGLRLRGKRAHGVRKGYTLRIIRA 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protist:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	39.3	273	16	051427
2	53	37.9	399	16	09K0X1
3	52	37.1	484	10	09SE92
4	52	37.1	1524	10	09S105
5	51	36.4	399	16	09J107
6	51	36.4	550	17	097W71
7	50	35.7	500	16	09A130
8	50	35.7	666	16	09I143
9	50	35.7	787	5	09NEL2
10	49	35.0	428	5	09TX80
11	49	35.0	545	17	0975A8
12	49	35.0	1689	5	09J135
13	49	35.0	1750	10	09SKJ4
14	49	35.0	2031	5	001306
15	49	35.0	2031	5	001307
16	49	35.0	2104	5	025440

17	49	35.0	2105	5	025439	025439 musca domes
18	49	35.0	2108	5	094615	094615 musca domes
19	48	34.3	132	10	09PVA4	09PVA4 tritricum ae
20	48	34.3	250	13	09IAJ7	09IAJ7 xenopus lae
21	48	34.3	556	16	097EB3	097EB3 clostiridium
22	48	34.3	633	17	09Y0U0	09Y0U0 pyrococcus
23	48	34.3	1972	10	09IUD7	09IUD7 arabidopsis
24	47.5	33.9	225	2	09EWE3	09EWE3 streptomyces
25	47	33.6	175	14	099IV8	099IV8 uncultured
26	47	33.6	258	5	09N4P4	09N4P4 caenorhabd1
27	47	33.6	519	4	09P2J1	09P2J1 homo sapien
28	47	33.6	520	16	09ZRC7	09ZRC7 rhizobium m
29	47	33.6	1053	16	09HY87	09HY87 pseudomonas
30	47	33.6	1951	4	09C007	09C007 homo sapien
31	47	33.6	1951	4	09EB23	09EB23 homo sapien
32	47	33.6	1984	11	008562	008562 rattus norv
33	47	33.6	2000	4	09C006	09C006 homo sapien
34	46	32.9	245	17	09HK58	09HK58 thermoplasm
35	46	32.9	279	11	054811	054811 cavia porce
36	46	32.9	283	2	0938V5	0938V5 bradyrhizob
37	46	32.9	445	16	09CNU0	09CNU0 pasteurella
38	46	32.9	450	11	09JUK2	09JUK2 mus musculu
39	46	32.9	492	16	09RZU9	09RZU9 delinococcus
40	46	32.9	510	11	062242	062242 mus musculu
41	46	32.9	515	16	0985I5	0985I5 rhizobium 1
42	46	32.9	521	16	098A10	098A10 rhizobium 1
43	46	32.9	1808	10	09LYS6	09LYS6 arabidopsis
44	46	32.9	1998	4	0961A3	0961A3 homo sapien
45	46	32.9	1999	4	09C008	09C008 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	273 AA.
051427	051427			
AC	051427:			
DT	01-JUN-1998 (TREMBLrel. 06, Created)			
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	CONSERVED HYPOTHETICAL PROTEIN.			
GN	BB0471.			
OS	Borrelia burgdorferi (Lyme disease spirochete).			
OC	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia.			
OX	NCBI_TaxID=139.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 35210 / B31;			
RX	MEDLINE=98065943; PubMed=9403685;			
RA	Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,			
RA	Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,			
RA	Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,			
RA	Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,			
RA	van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,			
RA	Ullrichback T., Wathey L., McDonald L., Arlrich P., Bowman C.,			
RA	Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,			
RA	Smith H.O., Venter J.C.,			
RT	"Genomic sequence of a Lyme disease spirochaete, Borrelia			
RT	burgdorferi."			
RL	Nature 390:580-586(1997).			
RL	EMBL: AB01151; AAC66832.1; -			
DR	TIGR: BB0471; -			
DR	InterPro: IPR003792; DUF189.			
DR	Pfam: PF02640; DUF189; 1.			
KW	Complete proteome.			
SQ	SEQUENCE 273 AA; 30683 MW; 7AEAI892E761A335 CRC64;			

Query Match 39.3%; Score 55; DB 16; Length 273;
Best Local Similarity 56.5%; Pred. No. 5.4;
Matches 13; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 GRLGKRIAGVKKYGYPTVL 24
 ||| ||| | : ||| | :
 DB 45 GRLGKRIAGVKKYGYPTVL 67

RESULT 2
 O9K0X1 PRELIMINARY: PRT: 399 AA.
 AC O9K0X1:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ACETATE KINASE.
 GN NMB0435.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MC58 / SEROGROUP B:
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tetteijn H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eelsen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwin M.L., Deboy R., Peterson J.D., Hickey E.K.,
 RA Hatt D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Ullerbach T.R., Knout H., Qin H., Vantheman J.,
 RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58".
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002400; AAP40873.1; -.
 DR TIGR: NMB0435; -.
 DR InterPro: IPR000890; Acetate_Kin.
 DR PRINTS: PR00471; ACETATE_KINASE.
 DR PROSITE: PS01075; ACETATE_KINASE_1; 1.
 DR PROSITE: PS01076; ACETATE_KINASE_2; 1.
 KW Kinase; Complete proteome.
 SO SEQUENCE 399 AA; 42795 MW; B672CDD77F8B7A78 CRC64;

Query Match 37.9%; Score 53; DB 16; Length 399;
 Best Local Similarity 45.0%; Pred. No. 16;
 Matches 9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 3 LRLGKRIAGVKKYGYPTVL 22
 : : | : ||| : | : ||
 DB 83 IKAIGHRIAGHGEKYSVYL 102

RESULT 3
 O9SF52 PRELIMINARY: PRT: 484 AA.
 AC O9SF52:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE NON-LTR REVERSE TRANSCRIPTASE.
 GN FL1F8_9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Maiti R., Rensing C.M., Koo H., Fujii C.Y., Ullerbach T.R.,
 RA Barnstead M.E., Bowman C.L., White O., Niernan W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome 3 BAC FL1F8 genomic sequence."
 SO Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC016661; AAF23283.1; -.
 KW RNA-directed DNA polymerase.
 SO SEQUENCE 484 AA; 55746 MW; E89EA732DBB9BDC9 CRC64;

Query Match 37.1%; Score 52; DB 10; Length 484;
 Best Local Similarity 55.6%; Pred. No. 27;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 LRLGKRIAGVKKYGYPT 20
 : || | : ||| : ||| |
 DB 443 IRRKGNKLAHLAKYGCCT 460

RESULT 4
 O9S105 PRELIMINARY: PRT: 1524 AA.
 AC O9S105:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE NON-LTR RETROELEMENT REVERSE TRANSCRIPTASE.
 GN AT2G31520.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Rensing C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana".
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC007071; AAD24831.1; -.
 DR InterPro: IPR000477; RVTSE.
 DR Pfam: PF00078; rvc; 1.
 KW RNA-directed DNA polymerase.
 SO SEQUENCE 1524 AA; 174648 MW; BDE2B81D6F72F4AA CRC64;

Query Match 37.1%; Score 52; DB 10; Length 1524;
 Best Local Similarity 55.6%; Pred. No. 92;
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 LRLGKRIAGVKKYGYPT 20
 : || | : ||| : ||| |
 DB 1483 IRRKGNKLAHLAKYGCCT 1500

RESULT 5
 O9JT07 PRELIMINARY: PRT: 399 AA.
 AC O9JT07:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ACETATE KINASE (EC 2.7.2.1).
 GN ACKR2 OR NMA2050.
 OS Neisseria meningitidis (serogroup A).

Query Match	Score 51	DB 16	Length 399
Best Local Similarity	44.4%	Pred. NO. 43	
Matches 12	Conservative 6	Mismatches 7	Indels 2
<p>1 KGLRLGKRIAHGK-KYG-PTVLRII 25</p> <p>1 : : : : 268 RAIKELGSIIVHGIRIKPKPTILIGV 294</p>			

Query Match	35.7%	Score 50	DB 16	Length 666
Query Match	35.7%	Score 50	DB 16	Length 500
Best Local Similarity	34.6%	Score No. 54		
Matches 9	Conservative 6	Mismatches 11	Indels 0	Gaps 0
QY 3	LRRLGRKIAHGKKYGYPTVLRIRIA 28			
DB 311	LAOLGNRIQIAFRAYTPKQKAVRVA 336			
RESULT 8				
Q91143	PRELIMINARY	PRT	666 AA	
AC Q91143	01-MAR-2001 (TREMblrel. 16, Created)			
DT 01-MAR-2001	(TREMblrel. 16, Last sequence update)			
DT 01-OCT-2001	(TREMblrel. 18, Last annotation update)			
DE	HYPOTHETICAL PROTEIN PA2439.			
GN PA2439.				
OS	Pseudomonas aeruginosa.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OC	Pseudomonas.			
OX	NCBI_TaxID=287;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 15692 / PA01;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.U., Brikanan F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Gatder R.L., Goltzy L., Tolentino E., Westbrock-Madman S., Yuan Y.,			
RA	Brody L.L., Coulter S.N., Folger K.R., Kas A., Lablign I.T.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,			
RA	Reiser J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PA01, an			
RT	opportunistic pathogen.";			
RL	Nature 406:959-964(2000).			
DR	EMBL; AE004671; AAG05827.1; -			
DR	InterPro; IPR001107; Band_7.			
DR	SMART; SMO0244; PHB; 1.			
KW	Hypothetical protein; Complete proteome.			
SQ	SEQUENCE 666 AA; 70532 MW; D30F396629015B69 CRC64;			

```

Best Local Similarity 50.0%; Pred. No. 74;
Matches 11; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 RGLRRLGRKIAHGKYGKPTVL 22
    | | | | | | | | | | | |
Db 118 RLRLRGLGAIAGLAQLGHAAL 139

RESULT 9
Q9NEL2 PRELIMINARY; PRT; 787 AA.
AC Q9NEL2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Y11B2A.22 PROTEIN.
GN Y11B2A.22.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peioderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Sulston J.E.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=8851916;
RA none;
RT "genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; AL132904; CAC35851.1;
SQ SEQUENCE 787 AA; 85108 MW; 4A1A9A55D03462B CRC64;

Query Match 35.7%; Score 50; DB 5; Length 787;
Best Local Similarity 37.0%; Pred. No. 88;
Matches 10; Conservative 7; Mismatches 8; Indels 2; Gaps 1;

QY 1 RGLRR--LGRKIAHGKYGKPTVLRII 25
    : | | | | | | | | | | | |
Db 691 QGIRRLPYGRPLRGVDNQAPTVAGV 717

RESULT 10
Q9TX80 PRELIMINARY; PRT; 428 AA.
AC Q9TX80;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SODIUM CHANNEL PROTEIN (FRAGMENT).
OS Musca domestica (House fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Diptera; Brachycera; Muscomorpha;
OC Muscidae; Muscidae; Musca.
OX NCBI_TaxID=7370;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93341454; PubMed=8101963;
RA Williamson M.S.; Denholm I.; Bell C.A.; Devonshire A.L.;
RT "knockdown resistance (kdr) to DDT and pyrethroid insecticides maps to
a sodium channel gene locus in the housefly (Musca domestica).";
RL Mol. Gen. Genet. 240:17-22(1993).
DR InterPro: IPR000636; Cation_chan_non_1lg.
DR InterPro: IPR002111; Cat_channel_TripL.
DR InterPro: IPR001682; Channel_pore_Ca_Na.
DR InterPro: IPR001696; Na_channel.
DR Pfam: PF00520; Ion_trans_1.
DR PRINTS; PR00170; NCHANNEL.
SQ SEQUENCE 428 AA; 46183 MW; 102985872C88931 CRC64;

```

```

Query Match 35.0%; Score 49; DB 5; Length 428;
Best Local Similarity 40.0%; Pred. No. 64;
Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1;

QY 6 LGRKIAHGKYY--GPTVLRIIRIA 28
    | | : : | | | | | | | | | |
Db 7 LGVLSDIIEKYVSPTLRVVRA 31

RESULT 11
Q975A8 PRELIMINARY; PRT; 545 AA.
AC Q975A8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE MOLYBDOPTERIN BIOSYNTHESIS MOEA PROTEIN.
GN ST0504.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000982; BAB65499.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 545 AA; 60835 MW; E21E5B8355F85C27 CRC64;

Query Match 35.0%; Score 49; DB 17; Length 545;
Best Local Similarity 46.2%; Pred. No. 83;
Matches 12; Conservative 6; Mismatches 6; Indels 2; Gaps 2;

QY 3 LRLRGRKIAHGK-KYG-PTVLRIIR 26
    : | | | | | | | | | | | |
Db 267 IRELGRKIIVHGKPKGKPTILATVK 292

RESULT 12
Q93135 PRELIMINARY; PRT; 1689 AA.
AC Q93135;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PARA-TYPE SODIUM CHANNEL (FRAGMENT).
OS Blattella germanica (German cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blattellidae; Blattellinae; Blattella.
OX NCBI_TaxID=6973;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C5MA;
RX MEDLINE=96397510; PubMed=8804404;
RA Miyazaki M., Ohyama K., Dunlap D.Y., Matsumura F.;
RT "Cloning and sequencing of the para-type sodium channel gene from
susceptible and kdr-resistant German cockroaches (Blattella germanica)
and house fly (Musca domestica).";
RL Mol. Gen. Genet. 252:61-68(1996).
DR EMBL; U71083; AAB82037.1;
DR InterPro: IPR000636; Cation_chan_non_1lg.

```

DR InterPro: IPR002111; Cat_channel_TrpL.
 DR InterPro: IPR002077; Ca_channel.
 DR InterPro: IPR001682; Channel_pore_Ca_Na.
 DR InterPro: IPR001064; Crystallin.
 DR InterPro: IPR001696; Na_channel.
 DR Pfam: PF00520; Ion_trans_4.
 DR PRINTS: PR00167; CCHANNEL.
 DR PRINTS: PR00170; NACHANNEL.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
 KM Ionic channel.1
 FT NON_TER 1
 FT VARIANT 785 785 .L->F.
 FT NON_TER 1689 1689
 SQ SEQUENCE 1689 AA; 191075 MW; EC8D12F826FAA028 CRC64;

Query Match 35.0%; Score 49; DB 5; Length 1669;
 Best Local Similarity 40.0%; Pred. No. 2.8e+02;
 Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1;

OY 6 LGRIAHGVKRY--GPTVLRIRIA 28
 DB 1456 LGVLSDIIEKYFVSPTLRYVRA 1480

RESULT 13
 O9SKJ4 PRELIMINARY; PRT: 1750 AA.
 AC O9SKJ4;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE NON-ITR RETROELEMENT REVERSE TRANSCRIPTASE.
 GN ATG325550.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_Taxid=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X.; Kaul S.; Rounsley S.D.; Shea T.P.; Benito M.-I.; Town C.D.;
 RA Fujii C.Y.; Mason T.M.; Bowman C.L.; Barnstead M.E.; Feldblum T.V.;
 RA Buell C.R.; Ketchum K.A.; Lee J.J.; Ronning C.M.; Koo H.; Moffat K.S.;
 RA Cronin L.A.; Shen M.; Vanden S.E.; Umayam B.; Talon L.J.; Gill J.E.;
 RA Adams M.D.; Carrera A.J.; Creasy T.H.; Goodman H.M.; Somerville C.R.;
 RA Copenhaver G.P.; Preuss D.; Nierman W.C.; White O.; Eisen J.A.;
 RA Salzberg S.L.; Fraser C.M.; Venter J.C.;
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana.";
 RT thaliana.402:761-768(1999).
 RL Mature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC006300; AAD20714.1;
 DR InterPro: IPR00477; RVTse.
 DR Pfam: PF00078; rvt.1.
 KM RNA-directed DNA polymerase.
 SQ SEQUENCE 1750 AA; 200736 MW; 03F504E81B37D3A6 CRC64;

Query Match 35.0%; Score 49; DB 10; Length 1750;
 Best Local Similarity 50.0%; Pred. No. 2.9e+02;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 OY 3 LRRLGRIAHGVKRYGPT 20
 DB 1709 IRKGNKLHVLAKYGC 1726

RESULT 14
 O01306 PRELIMINARY; PRT: 2031 AA.
 AC O01306;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PARA SODIUM CHANNEL.
 OS Blatella germanica (German cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberoidea; Blattellidae; Blattellinae; Blatella.
 NCBI_Taxid=6973;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C5MA;
 RX MEDLINE=97218696; PubMed=9066120;
 RA Dong K.;
 RT "A single amino acid change in the para sodium channel protein is
 associated with knockdown-resistance (kdr) to pyrethroid insecticides
 in German cockroach.";
 RL Insect Biochem. Mol. Biol. 27:93-100(1997).
 DR EMBL: U73583; AAC47483.1;
 DR InterPro: IPR000636; Catlon_chan_non_119.
 DR InterPro: IPR002111; Cat_channel_TrpL.
 DR InterPro: IPR001682; Channel_pore_Ca_Na.
 DR InterPro: IPR001064; Crystallin.
 DR InterPro: IPR001696; Na_channel.
 DR Pfam: PF00520; Ion_trans_4.
 DR PRINTS: PR00170; NACHANNEL.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
 KM Ionic channel.
 SQ SEQUENCE 2031 AA; 229962 MW; 228BAF57EEF5A54 CRC64;

Query Match 35.0%; Score 49; DB 5; Length 2031;
 Best Local Similarity 40.0%; Pred. No. 3.4e+02;
 Matches 10; Conservative 8; Mismatches 5; Indels 2; Gaps 1;

OY 6 LGRIAHGVKRY--GPTVLRIRIA 28
 DB 1664 LGVLSDIIEKYFVSPTLRYVRA 1688

RESULT 15
 O01307 PRELIMINARY; PRT: 2031 AA.
 AC O01307;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PARA SODIUM CHANNEL.
 OS Blatella germanica (German cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberoidea; Blattellidae; Blattellinae; Blatella.
 NCBI_Taxid=6973;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ECTIBAN-R;
 RX MEDLINE=97218696; PubMed=9066120;
 RA Dong K.;
 RT "A single amino acid change in the para sodium channel protein is
 associated with knockdown-resistance (kdr) to pyrethroid insecticides
 in German cockroach.";
 RL Insect Biochem. Mol. Biol. 27:93-100(1997).
 DR EMBL: U73584; AAC47484.1;
 DR InterPro: IPR000636; Catlon_chan_non_119.
 DR InterPro: IPR002111; Cat_channel_TrpL.
 DR InterPro: IPR001682; Channel_pore_Ca_Na.
 DR InterPro: IPR001064; Crystallin.
 DR InterPro: IPR001696; Na_channel.

Fri Jul 12 08:56:22 2002

us-09-642-744b-27.rspt

Page 6

```
DR Pfam: PF00520; Ion_trans_4.  
DR PRINTS; PRO0170; NACHANNEL.  
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.  
KW Ionic channel.  
SQ SEQUENCE 2031 AA; 229996 MW; 2F0821181E5FA6 CRC64;
```

Query Match	35.0%	Score 49	DB:5	Length 2031
Best Local Similarity	40.0%	Pred. No. 3	4e02	
Matches	10	Conservative	8	Mismatches 5, Indels 2, Gaps 1
OY	6	LGRKTAHGCVKKY--GPTVLRIRRA	28	
		:::		
Db	1664	LGVLVSDIIETKVFVSPTLTKVYKVA	1668	

Search completed: July 12, 2002, 08:03:56
Job time: 642 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:53:13 ; Search time 227.32 Seconds
(without alignments)
8.795 Million cell updates/sec

Title:	US-09-642-744B-1
Perfect score:	90
Sequence:	1 KNLRRIRKIHIHKKYG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

Database : A_Geneseq_032802:*

1	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1980.DAT*
2	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1981.DAT*
3	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1982.DAT*
4	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1983.DAT*
5	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1984.DAT*
6	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1985.DAT*
7	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1986.DAT*
8	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1987.DAT*
9	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1988.DAT*
10	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1989.DAT*
11	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1990.DAT*
12	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1991.DAT*
13	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1992.DAT*
14	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1993.DAT*
15	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1994.DAT*
16	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1995.DAT*
17	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1996.DAT*
18	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1997.DAT*
19	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1998.DAT*
20	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA1999.DAT*
21	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA2000.DAT*
22	/SDSI/gcgdata/hold_geneseq/geneseqp-emb1/AA2001.DAT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	100.0	18	22	AAB70648	Ovine SMAP 29 catch
2	90	100.0	18	22	AAB70655	Ovine SMAP 29 catch
3	90	100.0	29	22	AAB70675	Ovine SMAP 29 catch
4	88	97.8	18	22	AAB70654	Ovine SMAP 29 catch
5	88	97.8	20	22	AAB70653	Ovine SMAP 29 catch
6	79	87.8	16	22	AAB70656	Ovine SMAP 29 catch
7	77	85.6	16	22	AAB70657	Ovine SMAP 29 catch
8	66	73.3	14	22	AAB70658	Ovine SMAP 29 catch
9	66	73.3	14	22	AAB70659	Ovine SMAP 29 catch
10	64	71.1	14	22	AAB70660	Ovine SMAP 29 catch
11	64	71.1	14	22	AAB70661	Ovine SMAP 29 catch

12	62	8	9	13	22	AAB70663	Ovine SMAP 29 catch
13	61	67.8		12	22	AAB70662	Ovine SMAP 29 catch
14	57	63.3		12	22	AAB70664	Ovine SMAP 29 catch
15	57	63.3		18	22	AAB70652	Ovine SMAP 29 catch
16	57	63.3		28	22	AAB70674	Ovine SMAP 29 catch
17	57	63.3		29	22	AAB70649	Ovine SMAP 29 catch
18	45	50.0		224	21	AAB80789	M. jannaschii Mj08
19	45	50.0		1634	21	AAV52033	M. jannaschii Mj08
20	45	50.0		1634	21	AAV51652	M. jannaschii Mj08
21	43.5	48.3		167	22	AAV51652	S. epidermidis ope
22	43.5	48.3		208	21	AAV52204	Neisseria gonorrhe
23	43.5	48.3		208	21	AAV75364	Neisseria meningit
24	43.5	48.3		208	21	AAV75366	Neisseria meningit
25	43	47.8		20	22	AAB70651	Ovine SMAP 29 catch
26	43	47.8		291	22	AAG74953	Human colon cancer
27	43	47.8		1038	22	AAV52660	Novel human secret
28	42	46.7		187	19	AAV11090	H. pylori ORF hp5p
29	42	46.7		198	19	AAV11091	H. pylori ORF hp5p
30	42	46.7		447	22	ABG33413	Novel human diagen
31	42	46.7		460	22	ABH67825	Drosophila melanog
32	41.5	46.1		312	22	AAU33842	Staphylococcus aur
33	41.5	46.1		312	22	AAU36617	Staphylococcus aur
34	41.5	46.1		313	22	AAU37318	Staphylococcus aur
35	41.5	46.1		313	22	AAU37474	Staphylococcus aur
36	41	45.6		30	13	AAAR21399	Sequence of amphip
37	41	45.6		30	13	AAAR22883	Amphiphilic peptid
38	41	45.6		36	13	AAAR21400	Sequence of amphip
39	41	45.6		36	13	AAAR22884	Amphiphilic peptid
40	41	45.6		50	22	AAAB62000	B-myb protein f1rag
41	41	45.6		190	21	AAAG58947	Arabidopsis thailia
42	41	45.6		190	21	AAAG57222	Arabidopsis thailia
43	41	45.6		191	21	AAAG51404	Arabidopsis thailia
44	41	45.6		201	21	AAAG51721	Arabidopsis thailia
45	41	45.6		201	21	AAAG61403	Arabidopsis thailia

ALIGNMENTS

```
RESULT      1
AAB70648
ID   AAB70648 standard; peptide; 18 AA.
XX
AC   AAB70648;
XX
DT   15-MAY-2001 (first entry)
XX
DE   Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:1.
XX
KW   Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX   bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX   proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX   Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS   Ovis aries.
XX
PN   WO200112668-A1.
XX
PD   22-FEB-2001.
XX
PE   18-AUG-2000; 2000MO-US22781.
XX
PR   18-AUG-1999; 99US-0149886.
XX
PA   (IOWA ) UNIV IOWA RES FOUND.
XX   (REGC ) UNIV CALIFORNIA.
XX
PI   Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
DR   WPI; 2001-234911/24.
XX
PT   New antimicrobial peptides useful as antibiotics for inhibiting growth
XX   and proliferation of microbes, and for treating microbial infections
```

XX Claim 1; Page 103; 137pp; English.
PS
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 18 AA:

Query Match 100.0%; Score 90; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNLRRIRKIIHIKKYG 18
Db 1 knlrrlrkllhikkyg 18
|||||
|

RESULT 2
AAB70655 standard; peptide: 18 AA.
ID AAB70655 standard; peptide: 18 AA.
XX
AC AAB70655;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:8.
XX
XX
XX Ovine: SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KM *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX
XX Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
XX
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI
XX WPI: 2001-234911/24.
XX
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such

CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 18 AA:

Query Match 100.0%; Score 90; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNLRRIRKIIHIKKYG 18
Db 1 knlrrlrkllhikkyg 18
|||||
|

RESULT 3
AAB70675 standard; peptide: 29 AA.
ID AAB70675 standard; peptide: 29 AA.
XX
AC AAB70675;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:28.
XX
XX
XX Ovine: SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KM *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX
XX Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
XX
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI
XX WPI: 2001-234911/24.
XX
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 29 AA:

Query Match 100.0%; Score 90; DB 22; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNLRRIRKIIHIKKYG 18
|||||
|

DB 1 knlrrlrrkllhikkyg 18

RESULT 4
ID AAB70654 standard; peptide: 18 AA.
AC AAB70654;
XX
XX 15-MAY-2001 (first entry)
DT
XX
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:7.
XX
XX
KM Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
PD 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-US22781.
PF
XX 18-AUG-1999; 99US-0149886.
PR
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
DR WPI: 2001-234911/24.
XX
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
XX
PS Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
XX Sequence 18 AA;
SQ

Query Match 97.8%; Score 88; DB 22; Length 18;
Best Local Similarity 94.4%; Pred. No. 3.3e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNLRRRIIRKIIHIKKYG 18
DB 1 knlrrlrrkllhikkyg 18

RESULT 5
ID AAB70653 standard; peptide: 20 AA.
AC AAB70653;
XX
XX 15-MAY-2001 (first entry)
DT
XX
XX Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:6.

XX
XX Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
PD 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-US22781.
PF
XX 18-AUG-1999; 99US-0149886.
PR
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
DR WPI: 2001-234911/24.
XX
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
XX
PS Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
XX Sequence 20 AA;
SQ

Query Match 97.8%; Score 88; DB 22; Length 20;
Best Local Similarity 94.4%; Pred. No. 3.7e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNLRRRIIRKIIHIKKYG 18
DB 1 knlrrlrrkllhikkyg 18

RESULT 6
ID AAB70656 standard; peptide: 16 AA.
AC AAB70656;
XX
XX 15-MAY-2001 (first entry)
DT
XX
XX Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:9.
DE
XX
XX Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
PD 22-FEB-2001.
XX


```
PF 18-AUG-2000: 2000WO-US22781.
XX
PR 18-AUG-1999: 99US-0149886.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI: 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PS and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 1: Page 103: 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
SQ Sequence 16 AA:

Query Match 87.8%; Score 79; DB 22: Length 16;
Best Local Similarity 100.0%; Pred. No. 7.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NLRRIIRKIIHIHKY 17
|:|||||
Db 1 nlrriirkihikky 16

RESULT 7
AAB70657
ID AAB70657 standard: peptide: 16 AA.
XX
AC AAB70657;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:10.
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
XX WO200112668-A1.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000: 2000WO-US22781.
XX
XX 18-AUG-1999: 99US-0149886.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI: 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
```

```
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1: Page 103: 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
SQ Sequence 16 AA:

Query Match 85.6%; Score 77; DB 22: Length 16;
Best Local Similarity 93.8%; Pred. No. 1.6e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NLRRIIRKIIHIHKY 17
|:|||||
Db 1 nlrriirkihikky 16

RESULT 8
AAB70658
ID AAB70658 standard: peptide: 14 AA.
XX
XX AAB70658;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:11.
XX
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
XX WO200112668-A1.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000: 2000WO-US22781.
XX
XX 18-AUG-1999: 99US-0149886.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI: 2001-234911/24.
XX
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PS and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1: Page 103: 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
```

CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.

SO Sequence 14 AA:

Query Match 73.3%; Score 66; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.00072;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LRRIRKIIHIKK 16
 |||||
 Db 1 lrrirkihiikk 14

RESULT 9
 AAB70659
 ID AAB70659 standard; peptide; 14 AA.

AC AAB70659;

DT 15-MAY-2001 (first entry)

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:12.

XX Ovine: SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
 KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
 KW Burkholderia cepacia; Alcaligenes; Xanthomonas.

XX Ovis aries.

XX WO200112668-A1.

XX 22-FEB-2001.

PF 18-AUG-2000; 2000WO-US22781.

PR 18-AUG-1999; 99US-0149886.

XX (IOWA) UNIV IOWA RES FOUND.

PA (REGC) UNIV CALIFORNIA.

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX WPI: 2001-234911/24.

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (1)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.

SO Sequence 14 AA:

Query Match 73.3%; Score 66; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.00072;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LRRIRKIIHIKK 16

Db |||||
 1 lrrirkihiikk 14

RESULT 10

AAB70660
 ID AAB70660 standard; peptide; 14 AA.

AC AAB70660;

DT 15-MAY-2001 (first entry)

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:13.

XX Ovine: SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
 KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
 KW Burkholderia cepacia; Alcaligenes; Xanthomonas.

XX Ovis aries.

XX WO200112668-A1.

PD 22-FEB-2001.

PF 18-AUG-2000; 2000WO-US22781.

PR 18-AUG-1999; 99US-0149886.

XX (IOWA) UNIV IOWA RES FOUND.

PA (REGC) UNIV CALIFORNIA.

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX WPI: 2001-234911/24.

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (1)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.

SO Sequence 14 AA:

Query Match 71.1%; Score 64; DB 22; Length 14;
 Best Local Similarity 92.9%; Pred. No. 0.0015;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LRRIRKIIHIKK 16
 :|||
 Db 1 lrrirkihiikk 14

RESULT 11

AAB70661
 ID AAB70661 standard; peptide; 14 AA.

AC AAB70661;

DT 15-MAY-2001 (first entry)

XX

```

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:14.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
OS
XX Ovis aries.
XX
XX WO200112668-A1.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections
XX
XX PS Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX Xanthomonas.
XX
XX Sequence 14 AA:
SQ

```

```

XX
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections
XX
XX PS Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX Xanthomonas.
XX
XX Sequence 13 AA:
SQ

```

```

Query Match 68.9%; Score 62; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 RRIIRKIIHIKK 16
DB 1 RRIIRKIIHIKK 13

RESULT 13
AAB70662
ID AAB70662 standard; peptide: 13 AA.
XX
XX AAB70662:
XX
XX 15-MAY-2001 (first entry)
XX
XX Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:15.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
XX
XX WO200112668-A1.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX

```

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 1; Page 103; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 13 AA:

Query Match 67.8%; Score 61; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.004; Mismatches 0; Gaps 0;
Matches 13; Conservative 0; Indels 0;

QY 3 LRRIIRKIIHIIRK 15
Db 1 LRRIIRKIIHIIRK 13

RESULT 14
AAB70664
ID AAB70664 standard; peptide: 12 AA.
XX
AC AAB70664:
XX
DT 15-MAY-2001 (first entry)
XX
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:17.
XX
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI WPI; 2001-234911/24.
XX
DR
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX
PS Claim 1; Page 103; 137pp; English.
XX
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 13 AA:

CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 12 AA:

Query Match 63.3%; Score 57; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.016; Mismatches 0; Gaps 0;
Matches 12; Conservative 0; Indels 0;

QY 4 RRIIRKIIHIIRK 15
Db 1 RRIIRKIIHIIRK 12

RESULT 15
AAB70652
ID AAB70652 standard; peptide: 18 AA.
XX
AC AAB70652:
XX
DT 15-MAY-2001 (first entry)
XX
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:5.
XX
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI WPI; 2001-234911/24.
XX
DR
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX
PS Claim 10; Page 107; 137pp; English.
XX
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 18 AA:

Query Match 63.3%; Score 57; DB 22; Length 18;
Best Local Similarity 61.1%; Pred. No. 0.023; Mismatches 4; Indels 0;
Matches 11; Conservative 3; Mismatches 4; Indels 0;

Oy 1 KNLRIIRKIIHKKYG 18
: ||| | : |||
Db 1 rglrlgrklahgvkkyg 18

Search completed: July 12, 2002, 08:00:41
Job time: 448 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:53:13 : Search time 75.52 Seconds
(Without alignments)
5.822 Million cell updates/sec

Title: US-09-642-744B-1
Perfect score: 90
Sequence: 1 KNLRRIRKIIHIKKYG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/prodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	45.6	30	1	US-07-725-331-62
2	41	45.6	30	5	PCT-US91-05047-62
3	41	45.6	36	1	US-07-725-331-63
4	41	45.6	36	5	PCT-US91-05047-63
5	41	45.6	50	4	US-09-156-316-4
6	39.5	43.9	132	1	US-07-820-154A-2
7	39.5	43.9	132	1	US-07-820-154A-6
8	39.5	43.9	132	2	US-08-097-554A-2
9	39.5	43.9	132	2	US-08-097-554A-6
10	39.5	43.9	132	3	US-08-480-640A-2
11	39.5	43.9	132	3	US-08-480-640A-6
12	39.5	43.9	132	3	US-08-295-802-2
13	39.5	43.9	132	3	US-08-295-802-6
14	39.5	43.9	132	4	US-08-686-968C-102
15	39.5	43.9	132	4	US-08-488-237A-2
16	39.5	43.9	132	4	US-08-488-237A-6
17	39.5	43.9	132	4	US-08-375-992A-2
18	39.5	43.9	132	4	US-08-375-992A-6
19	39.5	43.9	132	5	PCT-US93-00324-2
20	39.5	43.9	132	5	PCT-US93-00324-6
21	39.5	43.9	677	3	US-08-480-640A-115
22	39.5	43.9	677	3	US-08-480-640A-193
23	39.5	43.9	677	3	US-08-295-802-115
24	39.5	43.9	677	4	US-08-686-968C-58
25	39.5	43.9	677	4	US-08-686-968C-193
26	39.5	43.9	677	4	US-08-488-237A-115
27	39.5	43.9	677	4	US-08-488-237A-193

28	39.5	43.9	677	4	US-08-375-992A-115	Sequence 115, App
29	39.5	43.9	677	4	US-08-375-992A-193	Sequence 193, App
30	39	43.3	114	1	US-08-031-399-3	Sequence 3, Appl1
31	39	43.3	114	1	US-08-031-399-6	Sequence 6, Appl1
32	39	43.3	114	1	US-08-031-399-12	Sequence 12, Appl1
33	39	43.3	114	1	US-08-393-305-3	Sequence 3, Appl1
34	39	43.3	114	1	US-08-393-305-6	Sequence 6, Appl1
35	39	43.3	114	1	US-08-726-817-3	Sequence 3, Appl1
36	39	43.3	114	1	US-08-726-817-6	Sequence 6, Appl1
37	39	43.3	114	1	US-08-504-042-3	Sequence 3, Appl1
38	39	43.3	114	1	US-08-504-042-6	Sequence 6, Appl1
39	39	43.3	114	1	US-08-504-042-12	Sequence 12, Appl1
40	39	43.3	114	2	US-08-725-969-3	Sequence 3, Appl1
41	39	43.3	114	2	US-08-725-969-6	Sequence 6, Appl1
42	39	43.3	114	2	US-08-794-524-3	Sequence 3, Appl1
43	39	43.3	114	2	US-08-794-524-6	Sequence 6, Appl1
44	39	43.3	114	4	US-09-189-193-3	Sequence 3, Appl1
45	39	43.3	114	4	US-09-189-193-6	Sequence 6, Appl1

ALIGNMENTS

```

RESULT 1
US-07-725-331-62
: Sequence 62, Application US/07725331
: Patent No. 5294605
:
GENERAL INFORMATION:
: APPLICANT: Houghten, Richard
: TITLE OF INVENTION: Amphiphilic Peptide Compositions and
: TITLE OF INVENTION: Analogues Thereof
: NUMBER OF SEQUENCES: 68
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dressler, Goldsmith, Sutter, Shore,
: ADDRESSEE: 6 Milnamow
: STREET: 180 No. 5294605th Stetson
: CITY: Chicago
: STATE: IL
: COUNTRY: USA
: ZIP: 60601
:
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.24
:
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/725,331
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/554,422
: FILING DATE: 19-JUL-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Gamsou, Edward P.
: REGISTRATION NUMBER: 29,381
: REFERENCE/DOCKET NUMBER: 421250-80
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 3126165418
: TELEFAX: 3126165460
: INFORMATION FOR SEQ ID NO: 62:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 30 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
:
FEATURE:
: OTHER INFORMATION: acetylated at N-terminus, may be
: OTHER INFORMATION: a C-terminal amide.
US-07-725-331-62

```



```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/05047
; FILING DATE: 19910717
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/554,422
; FILING DATE: 19-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: 421250-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3126165418
; TELEFAX: 3126165460
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: acetylated at N-terminus, may be
; OTHER INFORMATION: a C-terminal amide
PCT-US91-05047-63

```

```

Query Match          45.6%; Score 41; DB 5; Length 36;
Best Local Similarity 31.2%; Pred. No. 5.1;
Matches 5; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy 1 KNLRIIRKIIRK 16
    | | | | | | | | | |
Db 9 KLIKLIKLIKLIK 24

```

```

RESULT 5
US-09-156-316-4
; Sequence 4, Application US/09156316
; Patent No. 6183961
; GENERAL INFORMATION:
; APPLICANT: Bernstein, Harold S.
; APPLICANT: Coughlin, Shaun R.
; TITLE OF INVENTION: Methods and Compositions for Regulating Cell Cycle
; FILE REFERENCE: UCSF-020/01US
; CURRENT APPLICATION NUMBER: US/09/156,316
; CURRENT FILING DATE: 1998-09-18
; EARLIER APPLICATION NUMBER: 60/060,688
; EARLIER FILING DATE: 1997-09-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-156-316-4

```

```

Query Match          45.6%; Score 41; DB 4; Length 50;
Best Local Similarity 54.5%; Pred. No. 7;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 8 RKIIRIKRYG 18
    | | | | | | | | | |
Db 11 QKVIELVKRYG 21

```

```

RESULT 6
US-07-820-154A-2
; Sequence 2, Application US/07820154A

```

```

; Patent No. 5382425
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; APPLICANT: Junker M.S., David E
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,154A
; FILING DATE: 19920113
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-820-154A-2

```

```

Query Match          43.9%; Score 39.5; DB 1; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

```

```

Qy 1 KNLRIIRKIIRK 16
    | | | | | | | | | |
Db 8 KNAKVIKSIISLDIDIK 26

```

```

RESULT 7
US-07-820-154A-6
; Sequence 6, Application US/07820154A
; Patent No. 5382425
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; APPLICANT: Junker M.S., David E
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,154A
; FILING DATE: 19920113
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P

```


TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)977-9550
TELEFAX: (212)664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: AMINO ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Swinepox virus
STRAIN: Kasza
POSITION IN GENOME:
MAP POSITION: ~23.2
UNITS: %G
US-07-820-154A-6

Query Match 43.9%; Score 39.5; DB 1; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 1 KNRRIIRKIIHI--IKK 16
||:|:| |||: |||
Db 8 KNARKVSKIISLDIDIKK 26

RESULT 8
US-08-097-554A-2
Sequence 2, Application US/08097554A
Patent No. 5869312
GENERAL INFORMATION:
APPLICANT: Cochran Ph.D., Mark D
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 112
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,554A
FILING DATE: July 22, 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)977-9550
TELEFAX: (212)664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-097-554A-2

Query Match 43.9%; Score 39.5; DB 2; Length 132;

Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 1 KNRRIIRKIIHI--IKK 16
||:|:| |||: |||
Db 8 KNARKVSKIISLDIDIKK 26

RESULT 9
US-08-097-554A-6
Sequence 6, Application US/08097554A
Patent No. 5869312
GENERAL INFORMATION:
APPLICANT: Cochran Ph.D., Mark D
APPLICANT: Junker M.S., David E
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 112
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,554A
FILING DATE: July 22, 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)977-9550
TELEFAX: (212)664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Swinepox virus
STRAIN: Kasza
POSITION IN GENOME:
MAP POSITION: ~23.2
UNITS: %G
US-08-097-554A-6

Query Match 43.9%; Score 39.5; DB 2; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 1 KNRRIIRKIIHI--IKK 16
||:|:| |||: |||
Db 8 KNARKVSKIISLDIDIKK 26

RESULT 10
US-08-480-640A-2
Sequence 2, Application US/08480640A
Patent No. 6033904
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.

Query Match	43.98;	Score 39.5;	DB 3;	Length 132;
Best Local Similarity	52.68;	Pred. No. 30;		
Matches 10;	Conservative 3;	Mismatches 3;	Indels 3;	Gaps 1

QY 1 KNLRIIRKIIHI--IKK 16
 |||:|||||: |||
 Db 8 KNARKVISKIISLQDIKK 26

RESULT 13

US-08-295-802-6
 ; Sequence 6, Application US/08295802
 ; Patent No. 6127163
 ; GENERAL INFORMATION:
 ; APPLICANT: Cochran Ph.D., Mark D
 ; APPLICANT: Junker M.S., David E
 ; TITLE OF INVENTION: Recombinant Swinepox Virus
 ; NUMBER OF SEQUENCES: 188
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: John P. White
 ; STREET: 30 Rockefeller Plaza
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10112
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/295,802
 ; FILING DATE: Herewith
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212)977-9550
 ; TELEFAX: (212)664-0525
 ; TELEX: 422523
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 132 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; HYPOTHEICAL: YES
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 ; ORIGINAL SOURCE:
 ; ORGANISM: Swinepox virus
 ; STRAIN: Kasza
 ; POSITION IN GENOME:
 ; MAP POSITION: ~23.2
 ; UNITS: #6
 ; US-08-295-802-6

Query Match 43.9%; Score 39.5; DB 3; Length 132;
 Best Local Similarity 52.6%; Pred. No. 30;
 Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 1 KNLRIIRKIIHI--IKK 16
 |||:|||||: |||
 Db 8 KNARKVISKIISLQDIKK 26

RESULT 14

US-08-686-968C-102
 ; Sequence 102, Application US/08686968C
 ; Patent No. 6221361
 ; GENERAL INFORMATION:
 ; APPLICANT: Cochran, Mark D.
 ; APPLICANT: Junker, David E.
 ; TITLE OF INVENTION: Recombinant Swinepox Virus
 ; FILE REFERENCE: 39119-H/JML

; CURRENT APPLICATION NUMBER: US/08/686,968C
 ; CURRENT FILING DATE: 1986-07-25
 ; NUMBER OF SEQ ID NOS: 231
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 102
 ; LENGTH: 132
 ; TYPE: PR
 ; ORGANISM: Swinepox virus
 ; US-08-686-968C-102

Query Match 43.9%; Score 39.5; DB 4; Length 132;
 Best Local Similarity 52.6%; Pred. No. 30;
 Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 1 KNLRIIRKIIHI--IKK 16
 |||:|||||: |||
 Db 8 KNARKVISKIISLQDIKK 26

RESULT 15

US-08-488-237A-2
 ; Sequence 2, Application US/08488237A
 ; Patent No. 6251403
 ; GENERAL INFORMATION:
 ; APPLICANT: Cochran, Mark D.
 ; APPLICANT: Junker, David E.
 ; TITLE OF INVENTION: Recombinant Swinepox Virus
 ; NUMBER OF SEQUENCES: 225
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: John P. White
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/488,237A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P
 ; REGISTRATION NUMBER: 28,678
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 278-0400
 ; TELEFAX: (212) 391-0525
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 132 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-488-237A-2

Query Match 43.9%; Score 39.5; DB 4; Length 132;
 Best Local Similarity 52.6%; Pred. No. 30;
 Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 1 KNLRIIRKIIHI--IKK 16
 |||:|||||: |||
 Db 8 KNARKVISKIISLQDIKK 26

Search completed: July 12, 2002, 07:55:05
 Job time: 112 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:53:13 ; Search time 95.47 Seconds
(Without alignments)
18.117 Million cell updates/sec

Title: US-09-642-744B-1
Perfect score: 90
Sequence: 1 KNLRIIRKIIHIKKYG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	63.3	152	2	S68411
2	57	63.3	160	2	S68228
3	57	63.3	160	2	S68412
4	46	51.1	319	1	S19248
5	45.5	50.6	381	1	C64416
6	45	50.0	1634	2	E64410
7	44	48.9	1670	2	T06754
8	43.5	48.3	208	2	E81935
9	43.5	48.3	208	2	G81169
10	43	47.8	210	2	AC1577
11	43	47.8	234	2	I40822
12	43	47.8	234	2	D97108
13	43	47.8	420	2	S45630
14	43	47.8	537	2	C90389
15	42	46.7	168	2	T08831
16	42	46.7	186	2	B90116
17	42	46.7	187	2	H71819
18	42	46.7	187	2	H71819
19	42	46.7	187	2	H71819
20	42	46.7	187	2	H71819
21	42	46.7	187	2	H71819
22	42	46.7	187	2	H71819
23	42	46.7	187	2	H71819
24	42	46.7	187	2	H71819
25	42	46.7	187	2	H71819
26	42	46.7	187	2	H71819
27	42	46.7	187	2	H71819
28	42	46.7	187	2	H71819
29	42	46.7	187	2	H71819

30	42	46.7	937	2	T37241	olfactory channel
31	42	46.7	957	2	D88651	protein B0212.5 (1
32	42	46.7	1741	2	T15978	hypothetical prote
33	41.5	46.1	313	2	A89971	comp-binding-facto
34	41	45.6	117	2	T12722	hypothetical prote
35	41	45.6	181	2	B90246	TATA box binding p
36	41	45.6	190	2	A85360	hypothetical prote
37	41	45.6	198	2	S55311	TATA-binding prote
38	41	45.6	207	2	E70407	probable CDP-alcoh
39	41	45.6	256	2	A35340	H+-transporting AT
40	41	45.6	297	2	H70446	UTP--glucose-1-pho
41	41	45.6	337	2	S46010	hypothetical prote
42	41	45.6	338	2	D90017	hypothetical prote
43	41	45.6	367	2	S19172	cytochrome P450 2B
44	41	45.6	414	2	G64091	cell division prot
45	41	45.6	417	2	A33269	DNA primase (EC 2.

ALIGNMENTS

RESULT 1
S68411
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68411
R:Mahoney, M.M.; Lee, A.Y.; Bresinski-Caliguri, D.J.; Hutterer, K.M.
FEBS Lett. 377, 519-522, 1995
A:Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobia
A:Reference number: S68411; MUID:96140581
A:Accession: S68411
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-152 <MAH>
A:Cross-references: EMBL:X92757
C:Gene: SC5-2
C:Superfamily: cathelin; cystatin homology
F:14-122/Domain: signal sequence #status predicted <SIG>
F:14-122/Domain: cystatin homology <CYS>
F:21-123/Domain: propeptide #status predicted <PRO>
F:124-152/Product: cathelin-related protein 2 #status predicted <MAT>

Query Match 63.3% Score 57; DB 2: Length 152;
Best Local Similarity 61.1% Pred. No. 0.19;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNLRIIRKIIHIKKYG 18
: |||: ||| | : |||
Db 124 RGLRLRGKRIHGVKKYG 141

RESULT 2
S68228
M:Alternate names: cathelicidin
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68228
R:Bagella, L.; Scocchi, M.; Zanetti, M.
FEBS Lett. 376, 225-228, 1995
A:Title: cDNA sequences of three sheep myeloid cathelicidins.
A:Reference number: S68228; MUID:96105386
A:Accession: S68228
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-160 <BAG>
A:Cross-references: EMBL:L68854; NID:91161248; PIDN:AAA85470.1; PID:91161249
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-130/Domain: cystatin homology <CYS>

F:425-795/Cross-Link: peptide (Arg-Ser) #status predicted
F:882-1359/Cross-Link: peptide (Asn-Ser) #status predicted

Query Match 50.0%; Score 45; DB 2; Length 1634;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 KNRRIIRIKIHIIRIKY 17
DB 207 KNEKELIKITETLEKEY 223

RESULT 7

DNA-directed RNA polymerase I 190K chain homolog F15B8.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Mar-2000
C:Accession: T06754
R:Queletier, F.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoerge, W.; Salanoubat, M.; Mewes, submitted to the Protein Sequence Database, April 1999
A:Reference number: 215794
A:Accession: T06754
A:Molecule type: DNA
A:Residues: 1-1670 <Q>
A:Cross-References: EMBL:AL049660; GSPDB:GN00061; ATSP:F15B8.150
A:Experimental source: cultivar Columbia; BAC clone F15B8
C:Genetics:
A:Gene: ATSP:F15B8.150
A:Map position: 3
A:Intons: 21/3; 78/2; 131/3; 233/3; 306/3; 362/3; 417/1; 479/3; 567/3; 597/2; 646/3; 71
C:Superfamily: human DNA-directed RNA polymerase II largest chain

Query Match 48.9%; Score 44; DB 2; Length 1670;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 IIRKIHIIRIKY 18
DB 1569 IIRKIHVFKSYG 1581

RESULT 8

Probable phosphoribosylanthranilate isomerase (EC 5.3.1.24) NMA0890 [imported] - Neisseria
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: E81935
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, H.; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556
A:Accession: E81935
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-208 <P>
A:Cross-References: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84170.1; PID:g737960
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: trpF; NMA0890
C:Superfamily: phosphoribosylanthranilate isomerase; trpF homology
C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 48.3%; Score 43.5; DB 2; Length 208;
Best Local Similarity 47.4%; Pred. No. 28;
Matches 9; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

OY 1 KNRRIIRIKI-IHIIRIKY 18
DB 68 ONIRRIILAEVPIHIIOFHG 86

RESULT 9

N(5'-phosphoribosyl)anthranilate isomerase NMB0688 [imported] - Neisseria meningitidis
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: G81169
R:Rettelin, H.; Saunders, N.J.; Heideberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizzi, M.
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Science 287, 1809-1815, 2000
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: G81169
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-208 <P>
A:Cross-References: GB:AE002423; GB:AE002098; NID:g7225913; PIDN:AAF41106.1; PID:g722
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0688
C:Superfamily: phosphoribosylanthranilate isomerase; trpF homology

Query Match 48.3%; Score 43.5; DB 2; Length 208;
Best Local Similarity 47.4%; Pred. No. 28;
Matches 9; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

OY 1 KNRRIIRIKI-IHIIRIKY 18
DB 68 ONIRRIILAEVPIHIIOFHG 86

RESULT 10

precorrin isomerase homolog lln1156 [imported] - Listeria innocua (strain Cl1p11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AC1577
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec, D.; Dominguez-Bernal, G.; Duclaud, E.; Durand, L.; Dussurgey, O.; Entlian, K.D.; Fsihi, D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefel, J.; Kunh, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schluter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1577
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <G>
A:Cross-References: GB:AL592022; PIDN:CAC96387.1; PID:g16413615; GSPDB:GN00178
A:Experimental source: strain Cl1p11262
C:Genetics:
A:Gene: lln1156
C:Superfamily: Methanobacterium thermoautotrophicum precorrin isomerase

Query Match 47.8%; Score 43; DB 2; Length 210;
Best Local Similarity 53.8%; Pred. No. 33;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 6 IIRKIHIIRIKY 18
DB 62 VIOKIHVFNKNG 74

RESULT 11

sigK protein - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 21-Jul-2000
C:Accession: I40822
R:Saier, U.; Treuner, A.; Buchholz, M.; Santangelo, J.D.; Durre, P.
J. Bacteriol. 176, 6572-6582, 1994
A:Title: Sporulation and primary sigma factor homologous genes in *Clostridium acetobutylicum*
A:Reference number: I40609; MUID:95050216
A:Accession: I40822
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-234 <RES>
C:Cross-references: GB:L23317; NID:g2745925; PIDN:AAB94775.1; PID:g528974
C:Genetics:
A:Gene: sigK
A:Start codon: GTG
C:Superfamily: transcription initiation factor sigma K; transcription initiation factor F156-230/Domain: transcription initiation factor sigma katF homology <KTF>

Query Match 47.8%; Score 43; DB 2; Length 234;
Best Local Similarity 46.7%; Pred. No. 37;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 LRIIRKIHIHKY 17
: | : | : |||
Db 55 IERNLRVAHIKKY 69

RESULT 12
D97108
DNA-dependent RNA polymerase sigma chain [imported] - *Clostridium acetobutylicum*
C:Species: *Clostridium acetobutylicum*
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: D97108
R:Nolling, J.; Bleton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium*
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D97108
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK79655.1; PID:g15024653; GSPDB:GN00168
A:Experimental source: *Clostridium acetobutylicum* ATCC824
C:Genetics:
A:Gene: CAC1689
C:Superfamily: transcription initiation factor sigma K; transcription initiation factor

Query Match 47.8%; Score 43; DB 2; Length 234;
Best Local Similarity 46.7%; Pred. No. 37;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 LRIIRKIHIHKY 17
: | : | : |||
Db 55 IERNLRVAHIKKY 69

RESULT 13
S45630
DNA primase chain p48 - human
C:Species: *Homo sapiens* (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C:Accession: S45630
R:Stadlauer, F.; Brueckner, A.; Rehfuss, C.; Eckerskorn, C.; Lottspeich, F.; Foerster, J.; Blochem, 222, 781-793, 1994
A:Title: DNA replication in vitro by recombinant DNA-polymerase-alpha-primase.
A:Reference number: S45628; MUID:94298818
A:Accession: S45630
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-420 <STA>
C:Cross-references: EMBL:X74330; NID:g510405; PIDN:CAA52377.1; PID:g510406

```

C:Superfamily: DNA primase 50k chain

Query Match          47.8%; Score 43; DB 2; Length 420;
Best Local Similarity 52.9%; Pred. No. 64;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 KNLRRIRKRIHIIRKY 17
          : ||| |||||
Db      206 EKIHPRKSIIRIKKY 222

RESULT 14
C90389
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: C90389
R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: C90389
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1*537 <KUR>
A:Cross-references: GB:AE006641; NID:g13615498; PIRN:AAK42370.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO2200

```

```

Query Match          47.8%   Score 43;   DB 2;   Length 537;
Best Local Similarity 43.8%   Pred. No. 81;
Matches      7;   Conservative      4;   Indels      0;   Gaps      0;

OY      2  NLRRIIRKIHIIKKY  17
      ::  ||:  :|:|||||
Db      81  DNMTTKATLHLIKKY  96

RESULT  15
S77768
hypothetical protein MC037 - Mycoplasma capricolum (fragment)
C:Species: Mycoplasma capricolum
C:date: 09-Oct-1997 #sequence-revision 24-Oct-1997 #text-change 21-Jul-2000
C:Accession: S77768; S48590
R:Bork, P.; Ouzounis, C.; Casati, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert,
Mol. Microbiol. 16, 955-967, 1995
A>Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physi
A:Reference number: S77739; MUID:96059641
A:Accession: S77768
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-11 (BOR)
A:Cross-references: EMBL:Z33033; NID:9541693; PIDN:CAA83711.1; PID:94379130
A:Experimental source: ATCC 27343
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Genetic code: SGC3

```

Query	March	45.7%	Score 42;	DB 2;	Length 41;
Best	Local Similarity	50.0%	Pred. No. 10;		
Matches	8;	Conservative	4;	Mismatches	4;
				Indels	0;
				Gaps	0;
Qy	1	KNLRRIIRKIIHIIKK	16		
	:	:	:	:	:
Db	17	QELKEFIQMOTHIKK	32		

Fri Jul 12 08:55:51 2002

us-09-642-744b-1.rpr

Page 5

Search completed: July 12, 2002, 07:56:49
Job time: 216 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:53:48 ; Search time 46.15 Seconds
(Without alignments)
15.102 Million cell updates/sec

Title: US-09-642-744B-1
Perfect score: 90
Sequence: 1 KNLRIIRKIIHIKKYG 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	63.3	160	1	SC51_SHEEP
2	57	63.3	160	1	SC52_SHEEP
3	45.5	50.6	381	1	TH11_METJA
4	45	50.0	1634	1	DPOL_METJA
5	43	47.8	420	1	PR11_HUMAN
6	43	47.8	821	1	MCW6_HUMAN
7	42	46.7	93	1	Y435_METJA
8	42	46.7	152	1	ARGR_THEMA
9	42	46.7	507	1	MCW6_RAT
10	42	46.7	594	1	SYA_BORBU
11	42	46.7	670	1	REP_HAEIN
12	41	45.6	138	1	RR11_ARATH
13	41	45.6	159	1	MB28_BOVIN
14	41	45.6	198	1	TBP_SULSH
15	41	45.6	198	1	TBP_SULSO
16	41	45.6	256	1	ATPF_RAT
17	41	45.6	337	1	YB21_YEAST
18	41	45.6	414	1	FTSY_HAEIN
19	41	45.6	417	1	PR11_MOUSE
20	41	45.6	491	1	CPB1_RAT
21	41	45.6	491	1	CPB2_RAT
22	41	45.6	613	1	DEAD_HAEIN
23	41	45.6	686	1	MYBB_CHICK
24	41	45.6	700	1	MYBB_HUMAN
25	41	45.6	704	1	MYBB_MOUSE
26	40	44.4	105	1	Y795_PYRHO
27	40	44.4	189	1	TBP_THECE
28	40	44.4	197	1	TBP_SULAC
29	40	44.4	492	1	CPB3_RAT
30	40	44.4	552	1	YBUD_ECOLI
31	40	44.4	743	1	MYBB_XENLA
32	40	44.4	821	1	PSA_PYRHO
33	39.5	43.9	584	1	LMRA_LACLA

34	39.5	43.9	584	1	LMRA_LACLC	P97046 lactococcus
35	39	43.3	108	1	YE24_METJA	O58919 methanococc
36	39	43.3	156	1	Y330_RICPR	O92d18 rickettsia
37	39	43.3	162	1	IL15_CERAE	P40221 cercopithec
38	39	43.3	162	1	IL15_FELCA	O97687 felis silve
39	39	43.3	162	1	IL15_HUMAN	P40933 homo sapien
40	39	43.3	162	1	IL15_MACMU	P48092 macaca mula
41	39	43.3	162	1	IL15_PIG	O95253 sus scrofa
42	39	43.3	190	1	TBP_PYRKO	O52366 pyrococcus
43	39	43.3	304	1	IF2A_YEAST	P20459 saccharomyc
44	39	43.3	320	1	RT86_ECOLI	P23070 escherichia
45	39	43.3	358	1	Y074_METJA	O60380 methanococc

ALIGNMENTS

RESULT 1	ID	SC51_SHEEP	STANDARD:	PRT:	160 AA.
AC	P49928:				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 35, Last sequence update)				
DE	Cathelin-related peptide SC5 precursor 1 (antibacterial peptide SMAP-29)				
DE	29) (Myeloid antibacterial peptide SMAP-29).				
OS	Ovis aries (Sheep).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Caprinae; Ovis.				
OX	NCBI_TaxID:9940;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Bone marrow;				
RX	MEDLINE-96140581; Pubmed-8549789;				
RA	Maloney M.M., Lee A.Y., Brezinski-Caliguri D.J., Huttner K.M.;				
RT	"Molecular analysis of the sheep cathelin family reveals a novel				
RT	antimicrobial peptide.";				
RT	FEBS Lett. 377:519-522(1995).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RA	Huttner K.M., Maloney M.M.;				
RL	Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.				
CC	-1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.				
CC	-1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: X92757; CA63412.1; -				
DR	EMBL: U60600; AAB49715.1; -				
DR	InterPro: IPR001894; Cathelicidin.				
DR	Pfam: PF00666; Cathelicidins; 1.				
DR	Prodom: PD001838; Cathelicidin; 1.				
DR	PROSITE: PS00946; CATHELICIDINS_1; 1.				
DR	PROSITE: PS00947; CATHELICIDINS_2; 1.				
KW	Antibiotic; Signal.				
FT	SIGNAL	1	29		POTENTIAL.
FT	PROPEP	30	131		BY SIMILARITY.
FT	CHAIN	132	160		CATHELIN-RELATED PEPTIDE SC5.
FT	MOD_RES	30	30		PYRROLIDONE CARBOXYLIC ACID
FT					(BY SIMILARITY).
FT	DISULFID	86	97		BY SIMILARITY.
FT	DISULFID	108	125		BY SIMILARITY.
FT	SEQUENCE	160 AA;	17786 MW;		BD9B3859C432C249 CRC64;
SO					

Query Match 63.3%; Score 57; DB 1; Length 160;
 Best Local Similarity 61.1%; Pred. No. 0.041;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KNLRRIIRKIIHIKKYK 18
 : |||: ||| | : ||| |
 Db 132 RGLRRLGRKIIAHGVKKYG 149

RESULT 2
 SC52_SHEEP STANDARD; PRT; 160 AA.
 ID SC52_SHEEP
 AC P49929;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Cathelin-related peptide SC5 precursor 2 (Antibacterial peptide SMAP-29) (Myeloid antibacterial peptide SMAP-29).
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA MEDLINE=96140581; PubMed=8549789;
 RA Mahoney M.M., Lee A.Y., Brezinski-Caliguri D.J., Huttner K.M.;
 RT "Molecular analysis of the sheep cathelin family reveals a novel
 RT antimicrobial peptide."
 RL FEBS Lett. 377:519-522(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA MEDLINE=96105386; PubMed=7498547;
 RA Bagella L., Scocchi M., Zanetti M.;
 RT "cDNA sequences of three sheep myeloid cathelicidins."
 RL FEBS Lett. 376:225-228(1995).
 CC -1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.
 CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X92758; CAA63413.1; -;
 DR EMBL: L46854; AAA85470.1; -;
 DR InterPro: IPR001894; Cathelicidin.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00666; Cathelicidins. 1.
 DR ProDom: PD001838; Cathelicidin; 1.
 DR SMART: SM00043; C1; 1.
 DR PROSITE: PS00946; CATHELICIDINS_1; 1.
 DR PROSITE: PS00947; CATHELICIDINS_2; 1.
 DR Antibiotic; Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT PROPEP 30 131 BY SIMILARITY.
 FT CHAIN 132 160 CATHELIN-RELATED PEPTIDE SC5.
 FT MOD_RES 30 30 PYROGLUTAMINE CARBOXYLIC ACID
 FT (BY SIMILARITY).
 FT DISULFID 86 97 BY SIMILARITY.
 FT DISULFID 108 125 BY SIMILARITY.
 FT CONFLICT 28 28 S -> R (IN REF. 2).
 FT SEQUENCE 160 AA; 17742 MW; 4FB98A09355B51F CRC64;

Query Match 63.3%; Score 57; DB 1; Length 160;
 Best Local Similarity 61.1%; Pred. No. 0.041;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KNLRRIIRKIIHIKKYK 18
 : |||: ||| | : ||| |
 Db 132 RGLRRLGRKIIAHGVKKYG 149

RESULT 3
 TH11_METUA STANDARD; PRT; 381 AA.
 ID TH11_METUA
 AC Q58341;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable thiamine biosynthesis protein th11.
 GN TH11 OR M00931.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghagen N.S.M., Weidman J.D., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii."
 RL Science 273:1058-1073(1996).
 CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE THIAZOLE MOIETY (BY
 CC SIMILARITY).
 CC -1- PATHWAY: THIAMINE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO THE TH11 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U67536; AAB98933.1; -;
 DR TIGR: M00931; -;
 DR InterPro: IPR004114; THUMP.
 DR InterPro: IPR003720; TH11.
 DR Pfam: PF02568; Th11; 1.
 DR Pfam: PF02926; THUMP; 1.
 KW Thiamine biosynthesis; Complete proteome.
 SQ SEQUENCE 381 AA; 43436 MW; 0A31F1069DA3357B CRC64;

Query Match 50.6%; Score 45.5; DB 1; Length 381;
 Best Local Similarity 55.6%; Pred. No. 6;
 Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

OY 1 KNLRRIIRKIIHIKKYK 17
 : |||: ||| | : ||| |
 Db 27 KNLEELIRKNIKILRKY 44

RESULT 4
 DPOL_METUA STANDARD; PRT; 1634 AA.
 ID DPOL_METUA
 AC O58295;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) [Contains: Mja pol-1 intein; Mja pol-2
 DE intein].
 GN POL OR MJ0885.
 OS Methanococcus jannaschii.
 OC Archaea: Euryarchaeota: Methanococcales: Methanococcaceae:
 CC Methanococcus
 CC NCBI_TaxID=2190;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerecavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merriam J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Kleink H.-P., Fraser C.M., Hurt M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -i- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
 + (DNA)(N).
 CC -i- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
 CC -i- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL: U67532; AAB98889.1; -;
 DR HSSP: P56689; ITGO.
 DR TIGR: MJ0885; -;
 DR InterPro: IPR002064; DNA_pol_B.
 DR InterPro: IPR003586; HIntC.
 DR InterPro: IPR003587; HIntN.
 DR InterPro: IPR002203; InteIn.
 DR InterPro: IPR004042; InteIn_endonuc.
 DR Pfam: PF00136; DNA_pol_B_3
 DR Pfam: PF03104; DNA_pol_B_exo; 1.
 DR PRINTS: PR00106; DNAPOLB.
 DR SMART: SM00305; HIntC; 2.
 DR SMART: SM00306; HIntN; 2.
 DR SMART: SM00486; POLB; 1.
 DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 DR PROSITE: PS00881; PROTEIN_SPLICING; 2.
 DR Transferrase: DNA-directed DNA polymerase; DNA replication;
 KW DNA-binding; Autocatalytic cleavage; Protein splicing;
 KW Complete proteome.
 FT CHAIN 1 425 POL, 1ST PART (POTENTIAL).
 FT CHAIN 426 794 MJA_POL-1 INTEIN (POTENTIAL).
 FT CHAIN 795 882 POL, 2ND PART (POTENTIAL).
 FT CHAIN 883 1358 MJA_POL-2 INTEIN (POTENTIAL).
 FT CHAIN 1359 1634 POL, 3RD PART (POTENTIAL).
 FT SEQUENCE 1634 AA; 191708 MW; 84A1FAFAB1F97DDD CRC64;

Query Match 50.0%; Score 45; DB 1; Length 1634;
 Best Local Similarity 47.1%; Pred. No. 30;
 Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 KNLRIIRKIIRIKKY 17
 ID 207 KNEKLIIRKIIRIKKY 223

RESULT 5
 PRIL_HUMAN
 ID PRIL_HUMAN STANDARD: PRT: 420 AA.
 AC P49642.
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA primase small subunit (EC 2.7.7.-) (DNA primase 49 kDa subunit)
 DE (p49).
 GN PRIL1.
 OS Homo sapiens (Human).
 OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
 OC Mammalia: Eutheria: Primates: Catarrhini: Homiidae: Homo.
 CC NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94298818; PubMed=8026492;
 RA Stadlbauer F., Brueckner A., Rehness C., Eckerskorn C.,
 RA Lotspelsch F., Foerster V., Tseng B.Y., Nashauer H.P.;
 RT "DNA replication in vitro by recombinant DNA-polymerase-alpha-
 RT primase".;
 RL Eur. J. Biochem. 222:781-793(1994).
 RN [2]
 RP SEQUENCE OF 97-146 FROM N.A.
 RX MEDLINE=97422622; PubMed=9268648;
 RA Cloutier S., Hamel H., Champagne M., Yotov W.V.;
 RT "Mapping of the human DNA primase 1 (PRIM1) to chromosome 12q13.";
 RL Genomics 43:398-401(1997).
 CC -i- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
 RNA PRIMERS FOR THE OKAZAKI FRAGMENTS MADE DURING DISCONTINUOUS
 DNA REPLICATION.
 CC -i- SUBUNIT: HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT.
 CC -i- SIMILARITY: BELONGS TO THE EUKARYOTIC PRIMASE SMALL SUBUNIT
 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL: X74330; CA53277.1; -;
 DR EMBL: U89689; AAC51726.1; -;
 DR MIM: I76635; -;
 DR InterPro: IPR002755; DNA_primase_S.
 DR Pfam: PF01896; DNA_primase_S; 1.
 DR Transferrase: DNA replication; DNA-directed RNA polymerase; Primosome.
 KW ACT-SITE 44 44
 FT ACT-SITE 109 109 POTENTIAL.
 FT ACT-SITE 111 111 POTENTIAL.
 FT METAL 121 121 POTENTIAL.
 FT METAL 122 122 POTENTIAL.
 FT METAL 128 128 POTENTIAL.
 FT METAL 131 131 POTENTIAL.
 FT SEQUENCE 420 AA; 49902 MW; 9B5AC900EC3CCE8 CRC64;

Query Match 47.8%; Score 43; DB 1; Length 420;
 Best Local Similarity 52.9%; Pred. No. 16;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 KNLRIIRKIIRIKKY 17
 Db 206 EKTHPRIRKSTINIKKY 222

RESULT 6
 MCM6_HUMAN
 ID MCM6_HUMAN STANDARD: PRT: 821 AA.
 AC Q14566; Q13504; Q99859;

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA replication licensing factor MCM6 (P105MCM).
GN MCM6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE-97431351; PubMed-9286856;
RA Tsuruga H., Yabuta N., Hosoya S., Tamura K., Endo Y., Nojima H.;
RA "HSCMG": a new member of the human MCM/P1 family encodes a protein
RT homologous to fission yeast M155.
RL Genes Cells 2:381-399 (1997).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE-98184833; PubMed-9516426;
RA Holthoff H.P., Baack M., Richter A., Ritzl M., Knippers R.;
RA "Human protein MCM6 on HeLa cell chromatin."
RL J. Biol. Chem. 273:7320-7325 (1998).
RN 13
RP SEQUENCE OF 640-821 FROM N.A.
RX MEDLINE-97131582; PubMed-8977093;
RA Harvey C.B., Wang Y., Dermol D., Phillips A., Mantel N.,
RA Swallow D.M.;
RA "Characterization of a human homologue of a yeast cell division cycle
RT gene, MCM6, located adjacent to the 5' end of the lactase gene on
RL chromosome 2q21."
FEBS Lett. 398:135-135 (1996).
CC -1- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF A SINGLE ROUND OF DNA
CC REPLICATION DURING S PHASE. BINDS TO CHROMATIN DURING G1 AND
CC DETACH FROM IT DURING S PHASE AS IF IT LICENSES THE CHROMATIN TO
CC REPLICATE.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D84557; BAA12699.1; -
DR EMBL: U46838; AAC50766.1; -
DR EMBL: U67284; AAB48165.1; -
DR EMBL: U67281; AAB48165.1; JOINED.
DR EMBL: U67282; AAB48165.1; JOINED.
DR EMBL: U67283; AAB48165.1; JOINED.
DR MIM: 601806; -
DR InterPro: IPR001208; MCM.
DR Pfam: PF00493; MCM.1.
DR SMART: SM00350; MCM.1.
DR PROSITE: PS00847; MCM.1; 1.
DR PROSITE: PS50051; MCM.2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW DNA replication; ATP-binding; Cell cycle.
FT DOMAIN 346 553
FT NP_BIND 346 403 ATP (POTENTIAL).
FT CONFLICT 377 387 PRTGEGTSLR -> SKDNRRDLSS (IN REF. 2).
FT CONFLICT 495 495 A -> T (IN REF. 2).
FT CONFLICT 738 738 MISSING (IN REF. 3).
FT CONFLICT 790 790 L -> P (IN REF. 2).
SQ SEQUENCE 821 AA; 92889 MW; F94968EB25A3E501 CRC64;
Query Match 47.8%; Score 43; DB 1; Length 821;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 2 NLRRIKRIHIIRKY 17
| : ||| | : |
Db 768 NKRRIKVIHLRTHY 783

RESULT 7
Y435.METJA
ID Y435.METJA STANDARD; PRT; 93 AA.
AC 057877;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0435.
GN MJ0435.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MDLINE-96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sulten G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073 (1996).
CC -1- SIMILARITY: BELONGS TO THE M. JANNASCHII MJ0126 / MJ0128 / MJ0141 /
CC MJ0435 / MJ0604 / MJ1215 / MJ1305 / MJ1379 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67495; AAB98423.1; -
DR TIGR: MJ0435; -
DR InterPro: IPR002934; NTP_transf.
DR Pfam: PF01909; NTP_transf.2; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 93 AA; 10661 MW; AA6FD2014B942A68 CRC64;
Query Match 46.7%; Score 42; DB 1; Length 93;
Best Local Similarity 52.9%; Pred. No. 5.3;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RA Dimova D., Sakanyan V.;
RT "The autoregulated ArgR protein of the hyperthermophilic bacterium
RT Thermotoga neapolitana: thermostability and DNA binding properties.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RA MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
CC -1- FUNCTION: REGULATES ARGININE BIOSYNTHESIS GENES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: BELONGS TO THE ARGR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ132286; CAB65262.1; -
DR EMBL: AF001717; AAB35458.1; -
DR HSP: O31408; 1B4B.
DR TIGR: TM0371; -
DR InterPro: IPR001669; Arg-repress.
DR Pfam: PF01316; Arg-repressor; 1.
DR Pfam: PF02863; Arg-repressor; C.1.
DR ProDom: PD007402; Arg-repress; 1.
KW Transcription regulation; DNA-binding; Trans-acting factor; Repressor;
KW Complete proteome.
SQ SEQUENCE 152 AA; 17222 MW; BCD5384934D490E9 CRC64;

Query Match 46.7%; Score 42; DB 1; Length 152;
Best Local Similarity 36.7%; Pred. No. 8.6;
Matches 11; Conservative 3; Mismatches 4; Indels 12; Gaps 1;

OY 1 KNLRIIRKIIH-----LIKKYG 18
DB 6 KRROELIRKIIHEKKISNOFOIVEELKRYG 35

RESULT 9
MCM6_RAT STANDARD; PRT: 507 AA.
ID MCM6_RAT
AC 062724;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA replication licensing factor MCM6 (Intestinal DNA replication
DE protein) (Fragment).
GN MCM6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLTZMAN; TISSUE=Intestine;
RA MEDLINE=96011641; PubMed=7590274;
RA Sykes D.E., Weiser M.M.;

```

```

RT "Rat intestinal crypt-cell replication factor with homology to early
RT S-phase proteins required for cell division.";
RL Gene 163:243-247(1995).
RN [2]
RP REVISIONS.
RC STRAIN=HOLTZMAN;
RA Sykes D.E.;
RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF A SINGLE ROUND OF DNA
CC REPLICATION DURING S PHASE. BINDS TO CHROMATIN DURING G1 AND
CC DETACH FROM IT DURING S PHASE AS IF IT LICENSES THE CHROMATIN TO
CC REPLICATE.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U17565; AAC18424.1; -
DR InterPro: IPR001208; MCM.
DR Pfam: PF00493; MCM; 1.
DR SMART: SM00350; MCM; 1.
DR PROSITE: PS00847; MCM; 1.
DR PROSITE: PS50051; MCM; 2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW DNA replication; ATP-binding; Cell cycle.
FT NON_TER 1 1
FT DOMAIN 32 239 MCM.
FT NP_BIND 82 89 ATP (POTENTIAL).
SQ SEQUENCE 507 AA; 57369 MW; 03CE569FFCDD654 CRC64;

Query Match 46.7%; Score 42; DB 1; Length 507;
Best Local Similarity 43.8%; Pred. No. 28;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 NDLRIIRKIIHIKKY 17
DB 454 NKKRIEKVVRHRTHY 469

RESULT 10
SYA_BORBU STANDARD; PRT: 594 AA.
ID SYA_BORBU
AC 051238;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (Alars).
GN Alas OR BB0220.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RA MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,
RA Utterback T., Matthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi.";
RL Nature 390:580-586(1997).

```



```

CC -1- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC      diphosphate + L-alanyl-tRNA(Ala).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001132; AAC66604.1; -.
DR TIGR: BB0220; -.
DR InterPro: IPR002106; AA-trna_ligase_II.
DR InterPro: IPR002318; tRNA-syn_2c.
DR Pfam: PF01411; tRNA-syn_2c; 2.
DR PRINTS: PR00980; TRNASYNTHALA.
DR PROSITE: PS00179; AA-trna_ligase_II_1; FALSE_NEG.
DR PROSITE: PS00339; AA-trna_ligase_II_2; FALSE_NEG.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
DR Complete proteome.
DR KW SEQUENCE 594 AA: 67773 MW: 23378306DP02349A CXC64;

Query Match 46.7%; Score 42; DB 1; Length 594;
Best Local Similarity 56.2%; Pred. No. 33;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRRIRKIIHIHKYKG 18
Db 307 LRLIRSRIRYAKKLG 322
||||:|:|:|:|:|:|

RESULT 11
REP_HAEIN
ID REP_HAEIN STANDARD; PRT; 670 AA.
P44804;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ATP-dependent DNA helicase rep (EC 3.6.1.-).
GN REP OR H10649.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
CC NCBI_TaxID=727;
CX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; Pubmed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Keliavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McCreney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodok A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrman J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: REP HELICASE IS A SINGLE-STRANDED DNA-DEPENDENT
CC ATPASE INVOLVED IN DNA REPLICATION; IT CAN INITIATE UNWINDING
CC AT A NICK IN THE DNA. IT BINDS TO THE SINGLE-STRANDED DNA AND
CC ACTS IN A PROGRESSIVE FASHION ALONG THE DNA IN THE 3' TO 5'
CC DIRECTION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions in no way
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32748; AAC2309.1; -.
DR HSSP: P09980; IUD4.
DR TIGR: H10649; -.
DR InterPro: IPR000212; UVRD-helicase.
DR Pfam: PF00580; UVRD-helicase; 1.
DR Hydrolase: Helicase; DNA replication; ATP-binding; DNA-binding;
KM Complete proteome.
NP_BIND 22 ATP (POTENTIAL).
SQ SEQUENCE 670 AA; 77745 MW; 802A131251ED4E24 CRC64;

QY Query Match 46.7%; Score 42; DB 1; Length 670;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 RRIIRKIHIHIIKYG 18
1 1 1 1 1 1 1 1
Db 30 RVIINKIAHLIEKCG 44

RESULT 12
RRI1_ARATH STANDARD: PRT; 138 AA.
AC P56802:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chloroplast 30S ribosomal protein S11.
GN RPS11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RN RA SEQUENCE FROM N.A.
RP STRAIN=CV, COLUMBIA;
RX MEDLINE=20039611; PubMed=10574454;
RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.;
RT "Complete structure of the chloroplast genome of Arabidopsis
thaliana.";
RL DNA Res. 6:283-290(1999).
CC -I- SIMILARITY: BELONGS TO THE S11P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AP000423; BAA84418.1; -
CC InterPro: IPR001971; Ribosomal_S11.
CC Pfam: PF00411; Ribosomal_S11; 1.
CC Prodom: PD001010; Ribosomal_S11; 1.
CC ProSITE: PS00054; RIBOSOMAL_S11; 1.
KM Ribosomal protein; Chloroplast.
SQ SEQUENCE 138 AA; 15023 MW; 7A97E2B96CF7898 CRC64;

QY 1 KNLRIIRKIHIH 13

```

Db 20 KNVRIRIPKGVITHV 32

```

RESULT 13
ID MB28_BOVIN STANDARD; PRT; 159 AA.
AC PS4229;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Antibacterial peptide BMAP-28 precursor (Myeloid antibacterial peptide
  BMAP28).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
  OX Bovidae; Bovinae; Bos.
  NCBI_TaxID=9913;
  [1]
  SEQUENCE FROM N.A.
  RC TISSUE-Bone marrow;
  RX MEDLINE=97067059; PubMed=8910461;
  RA Skerlavaj B., Gennaro R., Bagella L., Merluzzi L., Risso A.,
  Zanetti M.;
  RA "Biological characterization of two novel cathelicidin-derived
  peptides and identification of structural requirements for their
  antimicrobial and cell lytic activities.";
  RT Eur. J. Biochem. 238:769-776(1996).
  RL J. Biol. Chem. 271:28375-28381(1996).
  [2]
  CHARACTERIZATION.
  RP MEDLINE=96300243; PubMed=8706679;
  RA Stofci P., Tossi A., Lemarcic B., Romeo D.;
  RA "Purification and structural characterization of bovine
  cathelicidins, precursors of antimicrobial peptides.";
  RT Eur. J. Biochem. 238:769-776(1996).
  CC -1- FUNCTION: EXERTS A POTENT ANTIMICROBIAL ACTIVITY AGAINST GRAM-
  NEGATIVE AND GRAM-POSITIVE BACTERIA, INCLUDING METHICILLIN-
  RESISTANT STAPHYLOCOCCUS AUREUS, AND FUNGI.
  CC -1- MASS SPECTROMETRY: MW=14414; MW_ERR=1; METHOD=Electrospray;
  CC RANGE=30-159.
  CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
  CC -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
  DR EMBL; X97609; CAA66208.1; -
  DR InterPro: IPR001894; Cathelicidin.
  DR Pfam: PF00666; Cathelicidins_1.
  DR ProDom: PD001838; Cathelicidin_1.
  DR PROSITE; PS00946; CATHELICIDINS_1; 1.
  DR PROSITE; PS00947; CATHELICIDINS_2; 1.
  KW Antibiotic; Signal.
  FT SIGNAL 1 29
  FT PROPEP 30 131
  FT CHAIN 132 159 ANTIBACTERIAL PEPTIDE BMAP-28.
  FT MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID.
  FT DISULFID 86 97 BY SIMILARITY.
  FT DISULFID 108 125 BY SIMILARITY.
  SQ SEQUENCE 159 AA; 17616 MW; 89B7CB46C5EBC367 CRC64;

```

Query Match 45.6%; Score 41; DB 1; Length 159;
 Best Local Similarity 56.2%; Pred. No. 13;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 3 LRRIRIKRIIHKRYG 18
 | : ||| : |||

Db 134 LRSGLRKILRAWKRYG 149

```

RESULT 14
ID TBP_SULSH STANDARD; PRT; 198 AA.
AC Q55031;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TATA-box binding protein (TATA-box factor) (TATA sequence-binding
  protein) (TBP) (Box A binding protein) (BAP).
OS Sulfolobus shibatae.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
  NCBI_TaxID=2286;
  [1]
  SEQUENCE FROM N.A. AND CHARACTERIZATION.
  RX MEDLINE=95303611; PubMed=7784182;
  RA Qureshi S.A., Baumann P.B., Rowlands T., Khoo B., Jackson S.P.;
  RT "Cloning and functional analysis of the TATA binding protein from
  Sulfolobus shibatae.";
  RL Nucleic Acids Res. 23:1775-1781(1995).
  CC -1- FUNCTION: GENERAL FACTOR THAT PLAYS A ROLE IN THE ACTIVATION OF
  CC ARCHAEAL GENES TRANSCRIBED BY RNA POLYMERASE. BINDS SPECIFICALLY
  CC TO THE TATA BOX PROMOTER ELEMENT WHICH LIES CLOSE TO THE POSITION
  CC OF TRANSCRIPTION INITIATION.
  CC -1- SIMILARITY: BELONGS TO THE TBP FAMILY.
  CC -----
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
  DR EMBL; U23419; AAC43403.1; -
  DR HSSP; Q57050; 1PCZ.
  DR InterPro: IPR000814; TFIID.
  DR Pfam: PF00352; TBP; 2
  DR PRINTS; PRO0686; TIFACTORID.
  DR PROSITE; PS00351; TFIID; 2.
  KW Transcription regulation; DNA-binding; Repeat.
  FT REPEAT 14 90
  FT REPEAT 105 181
  SQ SEQUENCE 198 AA; 22341 MW; 55CF391A7A7B163C3 CRC64;

```

Query Match 45.6%; Score 41; DB 1; Length 198;
 Best Local Similarity 38.9%; Pred. No. 16;
 Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 QY 1 KNLRIRIKRIIHKRYG 18
 | : : : ||| : |||

```

RESULT 15
ID TBP_SULSO STANDARD; PRT; 198 AA.
AC P58178;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TATA-box binding protein (TATA-box factor) (TATA sequence-binding
  protein) (TBP) (Box A binding protein) (BAP).
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
  NCBI_TaxID=2287;
  [1]
  SEQUENCE FROM N.A.

```

```

RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE-2132296; PubMed-11427726;
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweez M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -I- FUNCTION: GENERAL FACTOR THAT PLAYS A ROLE IN THE ACTIVATION OF
CC ARCHAEAL GENES TRANSCRIBED BY RNA POLYMERASE. BINDS SPECIFICALLY
CC TO THE TATA BOX PROMOTER ELEMENT WHICH LIES CLOSE TO THE POSITION
CC OF TRANSCRIPTION INITIATION.
CC -I- SIMILARITY: BELONGS TO THE TBP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE006715; AAK1225.1; ALT_INTT.
DR InterPro: IPR000814; TFIID.
DR Pfam: PF00352; TBP; 2.
DR PRINTS: PR00686; TIFACTORIID.
DR PROSITE: PS00351; TFIID; 2.
KM Transcription regulation: DNA-binding; Repeat; Complete proteome.
FT REPEAT 14 90 1.
FT REPEAT 105 181 2.
SQ SEQUENCE 198 AA; 22309 MW; 231448D0B2B026D7 CRC64;

```

```

Query Match          45.6%; Score 41; DB 1; Length 198;
Best Local Similarity 38.9%; Pred. No. 16;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 KNLRLIRKRIHIKKYG 18
   :|:::|:|
Db 77 EELIKAVKRIIKTKKYG 94

```

Search completed: July 12, 2002, 08:04:36
Job time: 648 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:53:14 ; Search time 174.7 Seconds
(Without alignments)
17.824 Million cell updates/sec

Title: US-09-642-744B-1
Perfect score: 90
Sequence: 1 KNLRIIRKIIHIKKYG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	60.0	146	2	09A1H3
2	50	55.6	255	10	09LJW2
3	48	53.3	500	3	09P6R0
4	46	51.1	319	2	005804
5	46	51.1	954	5	09G015
6	46	51.1	969	5	09U022
7	44	48.9	756	10	093XR9
8	44	48.9	840	12	093XR9
9	44	48.9	1670	10	09SVY0
10	43.5	48.3	208	16	09K0C6
11	43.5	48.3	208	16	09JVD1
12	43	47.8	210	16	092CL5
13	43	47.8	234	16	092CL5
14	43	47.8	379	2	09L7Z6
15	43	47.8	537	17	097WL6
16	43	47.8	4564	5	077075

17	42	46.7	41	2	048971	048971 mycoplasma
18	42	46.7	125	10	09AVZ1	09avz1 guillardia
19	42	46.7	146	16	098013	098013 mycoplasma
20	42	46.7	147	17	0976M0	0976m0 sulfolobus
21	42	46.7	168	10	096433	096433 glycine max
22	42	46.7	186	10	09AVW6	09avw6 guillardia
23	42	46.7	187	16	092JH9	092jh9 helicobacte
24	42	46.7	233	2	09L7X2	09l7x2 clostridium
25	42	46.7	234	5	0952R1	0952r1 caenorhabdi
26	42	46.7	275	10	09LRL7	09lrl7 arabidopsis
27	42	46.7	359	16	P73385	P73385 synechocyst
28	42	46.7	429	17	097VQ9	097vq9 sulfolobus
29	42	46.7	460	5	09VL03	09vl03 drosophila
30	42	46.7	483	17	P95941	P95941 sulfolobus
31	42	46.7	483	17	09C2F6	09c2f6 neurospora
32	42	46.7	540	5	091XZ6	091xz6 caenorhabdi
33	42	46.7	937	5	017469	017469 caenorhabdi
34	41.5	46.1	313	16	053606	053606 staphylococ
35	41.5	46.1	357	10	09XE30	09xe30 oryza sativ
36	41	45.6	104	2	045020	045020 borrelia bu
37	41	45.6	106	2	044868	044868 borrelia bu
38	41	45.6	117	9	080196	080196 methanobact
39	41	45.6	190	10	09S0H0	09s0h0 arabidopsis
40	41	45.6	207	16	067288	067288 aquifex aeo
41	41	45.6	217	1	0977R2	0977r2 uncultured
42	42	45.6	256	11	09C0Q7	09c0q7 mus musculu
43	41	45.6	290	2	09S0H6	09s0h6 borrelia bu
44	41	45.6	297	16	067602	067602 aquifex aeo
45	41	45.6	309	2	044823	044823 borrelia bu

ALIGNMENTS

RESULT 1
ID 09A1H3 PRELIMINARY; PRT; 146 AA.
AC 09A1H3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN L10.
GN RPL10.
OS Candidatus Carsonella ruddii.
OC Bacteria; Proteobacteria; gamma subdivision; Candidatus Carsonella.
OX NCBI_TaxID=114186;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=1125546; PubMed=11222582;
RA Clark M.A., Baumann L., Thao M.L., Moran N.A., Baumann P.;
RT "Degenerative Minimalism in the Genome of a Psyllid Endosymbiont.";
RL J. Bacteriol. 183:1853-1861(2001).
DR EMBL: AF274444; AAK17073.1; -;
SQ SEQUENCE 146 AA; 17749 MW; 4EF8D7D33EB17864 CRC64;

Query Match 60.0%; Score 54; DB 2; Length 146;
Best Local Similarity 55.6%; Pred. No. 1.2;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Oy 1 KNLRIIRKIIHIKKYG 18
Db 124 KNLRIIRKIIHIKKYG 141
RESULT 2
ID 09LJW2 PRELIMINARY; PRT; 255 AA.
AC 09LJW2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SIMILARITY TO SERINE/THREONINE KINASE.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLMBIA;
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLMBIA;
 RX MEDLINE=20363099; PubMed=10907853;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
 RT TAC and BAC clones.";
 RL DNA Res. 7:217-221(2000).
 DR EMBL: AP000388; BAB02946.1;
 DR InterPro: IPR002902; DUF26.
 DR Pfam: PF01657; DUF26; 2.
 KW Kinase.
 SO SEQUENCE 255 AA; 29435 MW; F75B3E52AEC418CE CRC64;

Query Match 55.6%; Score 50; DB 10; Length 255;
 Best Local Similarity 50.0%; Pred. No. 7.9;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 KNLRIIRKIIHIKKYG 18
 11::1111::11
 DB 48 KNLNOVIRNISHLRLXG 65

RESULT 3
 ID 09P6R0 PRELIMINARY; PRT; 500 AA.
 AC 09P6R0;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 15, Last annotation update)
 DE TRANSCRIPTION FACTOR IIB 70 KDA SUBUNIT.
 GN SPBC1387.10C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972H-;
 RA Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL354632; CAB89885.1;
 DR InterPro: IPR004366; Cyclin.
 DR InterPro: IPR000812; FTIB.
 DR Pfam: PF00382; transcript_fac2; 2.
 DR PRINTS: PR00685; TIFACTOR1IB.
 DR SMART: SM00385; CYCLIN; 2.
 SO SEQUENCE 500 AA; 56749 MW; 6CF1FDF9ADA2B1B2 CRC64;

Query Match 53.3%; Score 48; DB 3; Length 500;
 Best Local Similarity 42.9%; Pred. No. 30;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 2 NLRIIRKIIHIK 15
 1111111111111111
 DB 245 NFRRSYREVHVYK 258

RESULT 4 Query Match 51.1%; Score 46; DB 5; Length 954;
 Best Local Similarity 41.2%; Pred. No. 11e+02;

005804
 ID 005804 PRELIMINARY; PRT; 319 AA.
 AC 005804;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE RNA-DIRECTED DNA POLYMERASE FROM RETRON EC107 (EC 2.7.7.49)
 DE (REVERSE TRANSCRIPTASE).
 GN RT.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ECOR 70;
 RX MEDLINE=92204001; PubMed=1372675;
 RA Herzer P.J., Inouye S., Inouye M.;
 RT "Retron-Ec107 is inserted into the Escherichia coli genome by
 RT replacing a palindromic 34bp intergenic sequence.";
 RL Mol. Microbiol. 6:345-354(1992)
 CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE - N
 CC PYROPHOSPHATE + DNA(N) (EC 2.7.7.49).
 CC -1- MISCELLANEOUS: RETRONS MAY BE THE ANCESTORS OF RETROVIRUS.
 CC -1- SIMILARITY: TO OTHER REVERSE TRANSCRIPTASES IN BACTERIA AND
 CC RETROVIRUSES.
 DR EMBL: X62583; CAA44468.1;
 DR InterPro: IPR000477; RTase.
 DR Pfam: PF00078; rvt; 1.
 KW RNA-directed DNA polymerase; Nucleotidyltransferase;
 KW Transposable element.
 SO SEQUENCE 319 AA; 36363 MW; BA8E54DA76AC0FFC CRC64;

Query Match 51.1%; Score 46; DB 2; Length 319;
 Best Local Similarity 38.9%; Pred. No. 39;
 Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 KNLRIIRKIIHIKKYG 18
 1111111111111111
 DB 243 REARALROEVHLCKYK 260

RESULT 5
 ID 09G0I5 PRELIMINARY; PRT; 954 AA.
 AC 09G0I5;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ALANYL-TRNA SYNTHETASE.
 GN ALAS.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20542069; PubMed=11078517;
 RA Bunjun S., Stachopoulos C., Graham D., Min B., Kitabatake M.,
 RA Wang A.L., Wang C.C., Vivas C.P., Weiss L.M., Soll D.;
 RT "A dual-specificity aminoacyl-tRNA synthetase in the deep-rooted
 RT eukaryote Giardia lamblia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12997-13002(2000).
 DR EMBL: AF245445; AAG23137.1;
 DR InterPro: IPR002318; tRNA-synt_2c.
 DR Pfam: PF01411; tRNA-synt_2c; 2.
 DR PRINTS: PR00980; TRNASYNTHALA.
 KW Aminoacyl-tRNA synthetase.
 SO SEQUENCE 954 AA; 106800 MW; 2F39A8AFBE2C45B8 CRC64;

Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy 2 NLRRIIRKIIHIKKYG 18
 |||::||:|
 Db 368 NLRNVLRRVPHILVSRG 384

RESULT 6
 ID 090022 PRELIMINARY; PRT; 969 AA.
 AC 090022;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ALANL-TRNA SYNTHETASE (FRAGMENT).
 GN ALAS.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardiae; Giardia.
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chinnade J.W., Brown J.R., Schimmel P., Ribas de Pouplana L.;
 RT "Detection of an intermediate stage of Mitochondria Genesis";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF188719; AAF05594.1; -
 DR InterPro: IPR002318; tRNA-synt_2c.
 DR Pfam: PF01411; tRNA-synt_2c: 4.
 DR PRINTS: PR00980; TRNASYNTHALA.
 KW Aminoacyl-tRNA synthetase.
 FT NON_TER 969 969
 SQ SEQUENCE 969 AA; 108438 MW; 1A0BCB1EF3780C80 CRC64;

Query Match 51.1%; Score 46; DB 5; Length 969;
 Best Local Similarity 41.2%; Pred. No. 1.1e+02;
 Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy 2 NLRRIIRKIIHIKKYG 18
 |||::||:|
 Db 383 NLRNVLRRVPHILVSRG 399

RESULT 7
 ID 093XR9 PRELIMINARY; PRT; 756 AA.
 AC 093XR9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BG55 PROTEIN.
 GN BG55.
 OS Bruguiera gymnorhiza.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Rhizophoraceae; Bruguiera.
 OX NCBI_TaxID=39984;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA Barzai T., Hershkovits G., Katcoff D.J., Hanagata N., Dubinsky Z.,
 RA Karube I.;
 RT "Identification of mRNA transcripts differentially expressed in
 RT mangrove plant, Bruguiera gymnorhiza."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB061795; BAB55653.1; -
 SQ SEQUENCE 756 AA; 84482 MW; 1597F3D46C3E0809 CRC64;

Query Match 48.9%; Score 44; DB 10; Length 756;
 Best Local Similarity 40.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy 4 RRIIRKIIHIKKYG 18
 ::||:|
 Db 201 KKILLYLHLKKYG 215

RESULT 8
 ID 09DHN4 PRELIMINARY; PRT; 840 AA.
 AC 09DHN4;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 79R PROTEIN.
 GN 79R.
 OS Yaba-like disease virus (YLDV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Yatapoxvirus.
 OX NCBI_TaxID=132475;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lee H.J.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21176366; PubMed=11277691;
 RA Lee H.J., Essani K., Smith G.L.;
 RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus.";
 RL Virology 281:170-192(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Lee H.J.;
 RL Thesis (2000), Sir William Dunn School of Pathology, University of.
 DR EMBL: AJ293568; CAC21317.1; -
 SQ SEQUENCE 840 AA; 97985 MW; FAB9A0A5BE744491 CRC64;

Query Match 48.9%; Score 44; DB 12; Length 840;
 Best Local Similarity 52.9%; Pred. No. 1.9e+02;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 1 KNLRIIRKIIHIKKY 17
 ::||:|
 Db 334 ENLKEISEKIIFFKVKY 350

RESULT 9
 ID 09SVY0 PRELIMINARY; PRT; 1670 AA.
 AC 09SVY0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE I 190K CHAIN-LIKE PROTEIN (EC
 DE 2.7.7.6).
 GN F1588.150.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Benes V., Rechmann S., Borkova D., Ansgore W., Mewes H.W.,
 RA Mayer K.F.X., Lemke K., Schueller C., Quetier F., Salanoubat M.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 EU Arabidopsis sequencing project;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL049660; CAB41189.1; -
 DR InterPro: IPR000722; RNA_pol_A.
 DR InterPro: IPR002879; RNA_pol_A2.
 DR Pfam: PF00623; RNA_pol_A; 1.


```

OY 6 IIRKIIIRKYYG 18
DB 62 VIOKIIHLKNG 74

RESULT 13
OY 059264 PRELIMINARY: PRT: 234 AA.
AC 059264:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RNA POLYMERASE SIGMA FACTOR (DNA-DEPENDENT RNA POLYMERASE SIGMA
  SUBUNIT).
GN SICK OR CAC1689.
OS Clostridium acetobutylicum
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95050216; PubMed=7961408;
RA Sauer U., Treuner A., Buchholz M., Santangelo J.D., Durre P.;
RT "Sporulation and primary sigma factor homologous genes in Clostridium
  acetobutylicum."
RL J. Bacteriol. 176:6572-6582(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hilti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
  bacterium Clostridium acetobutylicum."
RL J. Bacteriol. 183:4823-4838(2001).
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
  ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
  THEN IS RELEASED.
CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
CC EMBL: L23317; AAB94775.1;
DR EMBL: AE007679; AAK79655.1;
DR HSSP: P00579; 1STG.
DR InterPro: IPR000943; Sigma70.
DR Pfam: PF00140; sigma70.1.
DR PROSITE: PS00715; SIGMA70.1; 1.
DR PROSITE: PS00716; SIGMA70.2; UNKNOWN.1.
KW DNA-binding; DNA-directed RNA polymerase; Sigma factor; Sporulation;
  Transcription regulation; Complete proteome.
SQ SEQUENCE 234 AA; 26586 MW; 091DD27A22BC1E4 CRC64;

Query Match 47.8%; Score 43; DB 16; Length 234;
Best Local Similarity 46.7%; Pred. No. 81;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 3 LRIIRKIIIRKYY 17
DB 55 IERNRLVIAHIVKY 69

RESULT 14
OY 091726 PRELIMINARY: PRT: 379 AA.
AC 091726:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE REPON REVERSE TRANSCRIPTASE.
GN RRT.
OS Salmonella enteritidis.
OX Plasmid low molecular weight plasmid I.

```

```

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella
OX NCBI_TaxID=592;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21189311; PubMed=11292805;
RA Rychlik I., Sedkova A., Gregorova D., Kapiskova R.;
RT "Low-Molecular-Weight Plasmid of Salmonella enterica Serovar
  Enteritidis Codes for Replicon Reverse Transcriptase and Influences
  Phage Resistance."
RL J. Bacteriol. 183:2852-2858(2001).
DR EMBL: AF218051; AAF72414.1;
DR InterPro: IPR000123; RNA_DNAPolys.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; RVT_1.
DR PRINTS: PR00866; RNADNAPOLMS.
KW Plasmid; RNA-directed DNA polymerase.
SQ SEQUENCE 379 AA; 43958 MW; 4A33CD21F35F9CDE CRC64;

```

```

Query Match 47.8%; Score 43; DB 2; Length 379;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 KNLRIIRKIIIRKYYG 18
DB 254 KENRRIYIVLYICKKG 271

RESULT 15
OY 097WL6 PRELIMINARY: PRT: 537 AA.
AC 097WL6:
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DE HYPOTHETICAL PROTEIN SSO2200.
GN SSO2200.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Polstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: AE006826; AAK42370.1;
DR InterPro: IPR002789; DUF87.
DR Pfam: PF01935; DUF87.1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 537 AA; 60140 MW; 752F08818FED5D57 CRC64;

```

```

Query Match 47.8%; Score 43; DB 17; Length 537;
Best Local Similarity 43.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 NLRRIIRKIIIRKYY 17
DB 81 DMNTIKALHLIKKY 96

```

Search completed: July 12, 2002, 08:03:44
Job time: 630 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:00:41 ; Search time 227.32 Seconds
(without alignments)
8.795 Million cell updates/sec

Title: US-09-642-744B-7
Perfect score: 90
Sequence: 1 KMLRLIRKIHIIKKYG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802: *
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT: *
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT: *
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT: *
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT: *
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT: *
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT: *
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT: *
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT: *
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT: *
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT: *
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT: *
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT: *
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT: *
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT: *
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT: *
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT: *
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT: *
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT: *
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT: *
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT: *
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT: *
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	18	22	AA70648
2	90	100.0	18	22	AA70655
3	90	100.0	29	22	AA70675
4	88	97.8	18	22	AA70654
5	88	97.8	20	22	AA70653
6	79	87.8	16	22	AA70656
7	77	85.6	16	22	AA70657
8	66	73.3	14	22	AA70658
9	66	73.3	14	22	AA70659
10	64	71.1	14	22	AA70660
11	64	71.1	14	22	AA70661

12	62	68.9	13	22	AA70663	Ovine SMAP 29 cath
13	61	67.8	13	22	AA70662	Ovine SMAP 29 cath
14	57	63.3	12	22	AA70664	Ovine SMAP 29 cath
15	57	63.3	18	22	AA70652	Ovine SMAP 29 cath
16	57	63.3	28	22	AA70674	Ovine SMAP 29 cath
17	57	63.3	29	22	AA70649	Ovine SMAP 29 cath
18	45	50.0	224	21	AAW90789	M. jannaschli M08
19	45	50.0	1634	21	AAV52023	M. jannaschli M08
20	45	50.0	1634	21	AAV51652	M. jannaschli M08
21	43.5	48.3	167	22	AAW82204	S. epidermidis ope
22	43.5	48.3	208	21	AAV75365	Neisseria gonorrhoe
23	43.5	48.3	208	21	AAV75365	Neisseria meningit
24	43.5	48.3	208	22	AAW70651	Ovine SMAP 29 cath
25	43	47.8	20	22	AAW70651	Ovine SMAP 29 cath
26	43	47.8	291	22	AAW74583	Human colon cancer
27	43	47.8	1038	22	AAU33842	Human colon cancer
28	42	46.7	187	19	AAV11090	H. pylori ORF hp6p
29	42	46.7	198	19	AAV11091	H. pylori ORF hp6p
30	42	46.7	447	22	ABG23413	Novel human diago
31	42	46.7	460	22	ABB67825	Drosophila melanog
32	41.5	46.1	312	22	AAU33842	Staphylococcus aur
33	41.5	46.1	313	22	AAU36617	Staphylococcus aur
34	41.5	46.1	313	22	AAU37318	Staphylococcus aur
35	41.5	46.1	313	22	AAU37474	Staphylococcus aur
36	41	45.6	30	13	AAW21399	Sequence of amphip
37	41	45.6	30	13	AAW22883	Amphiphilic peptid
38	41	45.6	36	13	AAW21400	Sequence of amphip
39	41	45.6	36	13	AAW22884	Amphiphilic peptid
40	41	45.6	50	22	AAW62000	B-myb protein frag
41	41	45.6	190	21	AAW58947	Arabidopsis thalia
42	41	45.6	191	21	AAW57222	Arabidopsis thalia
43	41	45.6	191	21	AAW61404	Arabidopsis thalia
44	41	45.6	201	21	AAW57221	Arabidopsis thalia
45	41	45.6	201	21	AAW61403	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAB70648
AAB70648 standard; peptide: 18 AA.
ID AAB70648;
XX
AC AAB70648;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:1.
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
RD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
XX (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehner R;
XX WPI; 2001-234911/24.
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
and proliferation of microbes, and for treating microbial infections -

XX Claim 1; Page 103; 137pp; English.
 PS
 CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SNAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.
 CC
 XX Sequence 18 AA:
 SQ
 Query Match 100.0%; Score 90; DB 22; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1.0e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KNLRRIRKRIHIIRKYG 18
 Db 1 KNLIRIRKIHIIKKYG 18
 RESULT 2
 AAB70655
 ID AAB70655 standard; peptide: 18 AA.
 AC AAB70655:
 XX 15-MAY-2001 (first entry)
 DT
 XX Ovine SNAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:8.
 DE
 XX Ovine; SNAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
 KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
 KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
 XX
 OS Ovis aries.
 XX
 PN WO200112668-A1.
 PD 22-FEB-2001.
 PT
 XX New antimicrobial peptides useful as antibiotics for inhibiting growth
 XX and proliferation of microbes, and for treating microbial infections -
 XX
 PF 18-AUG-2000; 2000WO-US22781.
 PR 18-AUG-1999; 99US-0149886.
 PA (IOWA) UNIV IOWA RES FOUND.
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
 XX WPI: 2001-234911/24.
 DR
 XX New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 XX Claim 1; Page 103; 137pp; English.
 PS
 CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SNAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.

CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.
 CC
 XX Sequence 18 AA:
 SQ
 Query Match 100.0%; Score 90; DB 22; Length 18;
 Best Local Similarity 100.0%; Pred. No. 1.0e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KNLRRIRKRIHIIRKYG 18
 Db 1 KNLIRIRKIHIIKKYG 18
 RESULT 3
 AAB70675
 ID AAB70675 standard; peptide: 29 AA.
 AC AAB70675:
 XX 15-MAY-2001 (first entry)
 DT
 XX Ovine SNAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:28.
 DE
 XX Ovine; SNAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
 KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
 KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
 XX
 OS Ovis aries.
 XX
 PN WO200112668-A1.
 PD 22-FEB-2001.
 PT
 XX New antimicrobial peptides useful as antibiotics for inhibiting growth
 XX and proliferation of microbes, and for treating microbial infections -
 XX
 PF 18-AUG-2000; 2000WO-US22781.
 PR 18-AUG-1999; 99US-0149886.
 PA (IOWA) UNIV IOWA RES FOUND.
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
 XX WPI: 2001-234911/24.
 DR
 XX New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 XX Claim 1; Page 103; 137pp; English.
 PS
 CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SNAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.
 CC
 XX Sequence 29 AA:
 SQ
 Query Match 100.0%; Score 90; DB 22; Length 29;
 Best Local Similarity 100.0%; Pred. No. 2.6e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KNLRRIRKRIHIIRKYG 18
 Db 1 KNLIRIRKIHIIKKYG 18

XX	Ovine, SMAP29; lupine, RCAP 18; cathelicidin; antimicrobial;	
KM	bactericidal; antibiotic; antiviral; microbial growth inhibitor;	
KW	proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;	
XX	Burkholderia cepacia; Alcaligenes; Xanthomonas.	
OS	Ovis aries.	
PN	WO200112668-A1.	
PD	22-FEB-2001.	
XX		
PF	18-AUG-2000; 2000WO-US22781.	
XX		
PR	18-AUG-1999; 99US-0149886.	
XX		
PA	(IOWA) UNIV IOWA RES FOUND.	
XX	(REGC) UNIV CALIFORNIA.	
PI	Tack BF, McCray P, Welsh M, Travis SM, Lehrer R;	
XX		
DR	WPI; 2001-234911/24.	
PT		
XX	New antimicrobial peptides useful as antibiotics for inhibiting growth	
XX	and proliferation of microbes, and for treating microbial infections -	
XX		
PS	Claim 1; Page 103; 137pp; English.	
XX		
CC	AAB70648 to AAB70675 represent antimicrobial peptides (1), of which	
CC	AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine	
CC	SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are	
CC	derived from the lupine RCAP 18 cathelicidin family peptide. (1) have	
CC	antibiotic, antimicrobial and antiviral activities, and can be used as	
CC	microbial growth and proliferation inhibitors and in gene therapy. (1)	
CC	are useful for inhibiting microbial growth in an environment capable of	
CC	sustaining such growth, for inhibiting microbial growth or strain in a	
CC	host, and inhibiting the growth of drug-resistant microbial strains such	
CC	as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and	
CC	Xanthomonas.	
XX		
SQ	Sequence 20 AA:	
	Query Match 97.8%, Score 88; DB 22; Length 20;	
	Best Local Similarity 94.4%; Pred. NO. 3.7e-07;	
	Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
OY	1 KNLRIIRKRIHIIRKKYG 18	
	:	
DB	1 KNIRIRKRIHIIRKKYG 18	
RESULT 6		
AAB70656		
ID	AAB70656 standard; peptide: 16 AA.	
AC	AAB70656;	
XX		
DT	15-MAY-2001 (first entry)	
XX		
DE	Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:9.	
XX		
KW	Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;	
KW	bactericidal; antibiotic; antiviral; microbial growth inhibitor;	
KW	proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;	
KW	Burkholderia cepacia; Alcaligenes; Xanthomonas.	
XX		
OS	Ovis aries.	
XX		
PN	WO200112668-A1.	
XX		
PD	22-FEB-2001.	

```

PT and proliferation of microbes, and for treating microbial infections
XX
XX
XX Claim 1; Page 103; 137pp; English.
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
SQ Sequence 16 AA:
    Query Match 85.6% Score 77; DB 22; Length 16;
    Best Local Similarity 93.8% Pred. No. 1.6e-05;
    Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 NLRRIIRKIHIHKY 17
   1:|||||
Db 1 NLRRIIRKIHIHKY 16

```

XX	
XX	AAB70658;
AC	
XX	
DT	15-MAY-2001 (first entry)
XX	
DE	Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:11.
XX	
XX	Ovine: SMAP29: lupine; RCAP 18: cathelicidin; antimicrobial;
KW	bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW	proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW	Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX	
OS	Ovis aries.
XX	
PN	WO200112668-A1.
XX	
PD	22-FEB-2001.
XX	
PF	18-AUG-2000; 2000WO-US22781.
XX	
PR	18-AUG-1999; 99US-0149886.
XX	
PA	(IOWA) UNIV IOWA RES FOUND.
PA	(RESC) UNIV CALIFORNIA.
XX	
PI	Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX	
DR	WPI: 2001-234911/24.
XX	
PT	New antimicrobial peptides useful as antibiotics for inhibiting growth
PT	and proliferation of microbes, and for treating microbial infections -
XX	
PS	Claim 1; Page 103; 137pp; English.
XX	
XX	AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC	AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC	SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC	derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC	antibiotic, antimicrobial and antiviral activities, and can be used as
CC	microbial growth and proliferation inhibitors and in gene therapy. (I)
CC	are useful for inhibiting microbial growth in an environment capable of
CC	sustaining such growth, for inhibiting microbial growth or strain in a
CC	

Query Match	Best Local Similarity	Score 64;	DB 22;	Length 14;
Matches 13;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
OY 3 LRRIRKIIHIKK 16	1 LRRIRKIIHIKK 14			
<p>RESULT 12</p> <p>ID AAB70663</p> <p>XX AAB70663 standard; peptide: 13 AA.</p> <p>XX AAB70663;</p> <p>XX 15-MAY-2001 (first entry)</p> <p>DE ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:16.</p> <p>XX</p> <p>KW Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;</p> <p>KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;</p> <p>KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;</p> <p>KW Burkholderia cepacia; Alcaligenes; Xanthomonas.</p> <p>XX</p> <p>OS Ovis aries.</p> <p>XX</p> <p>XX WO200112668-A1.</p> <p>XX 22-FEB-2001.</p> <p>PD 18-AUG-2000; 2000WO-US22781.</p> <p>XX</p> <p>PR 18-AUG-1999; 99US-0149886.</p> <p>XX</p> <p>PA (IOWA) UNIV IOWA RES FOUND.</p> <p>PA (REGC) UNIV CALIFORNIA.</p> <p>XX</p> <p>PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;</p> <p>XX</p> <p>DR WPI: 2001-234911/24.</p> <p>FT New antimicrobial peptides useful as antibiotics for inhibiting growth</p> <p>PT and proliferation of microbes, and for treating microbial infections</p> <p>XX</p> <p>PS Claim 1; Page 103; 137pp; English.</p> <p>XX</p> <p>CC AAB70668 to AAB70675 represent antimicrobial peptides (1), of which</p> <p>CC AAB70668 to AAB70664, AAB70674 and AAB70675 are derived from the ovine</p> <p>CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are</p> <p>CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have</p> <p>CC antibiotic, antimicrobial and antiviral activities, and can be used as</p> <p>CC microbial growth and proliferation inhibitors and in gene therapy. (1)</p> <p>CC are useful for inhibiting microbial growth in an environment capable of</p> <p>CC sustaining such growth, for inhibiting microbial growth or strain in a</p> <p>CC host, and inhibiting the growth of drug-resistant microbial strains such</p> <p>CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and</p> <p>CC Xanthomonas.</p> <p>XX</p> <p>SO Sequence 14 AA.</p>				

[illegible]

PR New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 1; Page 103; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 13 AA;

Query Match 67.8%; Score 61; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LRRIRKIIHIK 15
DB 1 LRRIRKIIHIK 13

RESULT 14
AAB70664
ID AAB70664 standard; peptide; 12 AA.
XX
AC AAB70664;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:17.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KM *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX Ovis aries.
OS
XX
XX WO200112668-A1.
PN
XX
XX 22-FEB-2001.
PD
XX
XX 18-AUG-2000; 2000WO-US22781.
PF
XX
XX 18-AUG-1999; 99US-0149886.
PR
XX
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI
XX
XX WPI; 2001-234911/24.
DR
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ

CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 12 AA;

Query Match 63.3%; Score 57; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 RRIIRKIIHIK 15
DB 1 RRIIRKIIHIK 12

RESULT 15
AAB70652
ID AAB70652 standard; peptide; 18 AA.
XX
AC AAB70652;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:5.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KM *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX Ovis aries.
OS
XX
XX WO200112668-A1.
PN
XX
XX 22-FEB-2001.
PD
XX
XX 18-AUG-2000; 2000WO-US22781.
PF
XX
XX 18-AUG-1999; 99US-0149886.
PR
XX
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI
XX
XX WPI; 2001-234911/24.
DR
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 10; Page 107; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 18 AA;

Query Match 63.3%; Score 57; DB 22; Length 18;
Best Local Similarity 61.1%; Pred. No. 0.023;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 KNLRIIRKIHITKKY 18
: |||: ||| | : ||| |
Db 1 rglrIgrKiahgvKKYg 18

Search completed: July 12, 2002, 08:00:41
Job time: 448 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:55:05 ; Search time 75.52 Seconds
(without alignments)
5.822 Million cell updates/sec

Title: US-09-642-744B-7
Perfect score: 90
Sequence: 1 KMLRRIRKIHIIKKYG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCPUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	41	45.6	30	1	US-07-725-331-62
2	41	45.6	30	5	PCT-US91-05047-62
3	41	45.6	36	1	US-07-725-331-63
4	41	45.6	36	5	PCT-US91-05047-63
5	41	45.6	50	4	US-09-156-316-4
6	39.5	43.9	132	1	US-07-820-154A-2
7	39.5	43.9	132	1	US-07-820-154A-6
8	39.5	43.9	132	2	US-08-097-554A-2
9	39.5	43.9	132	2	US-08-097-554A-6
10	39.5	43.9	132	3	US-08-480-640A-2
11	39.5	43.9	132	3	US-08-480-640A-6
12	39.5	43.9	132	3	US-08-295-802-2
13	39.5	43.9	132	3	US-08-295-802-6
14	39.5	43.9	132	4	US-08-686-968C-102
15	39.5	43.9	132	4	US-08-488-237A-2
16	39.5	43.9	132	4	US-08-488-237A-6
17	39.5	43.9	132	4	US-08-375-992A-2
18	39.5	43.9	132	4	US-08-375-992A-6
19	39.5	43.9	132	5	PCT-US93-00324-2
20	39.5	43.9	132	5	PCT-US93-00324-6
21	39.5	43.9	677	3	US-08-480-640A-115
22	39.5	43.9	677	3	US-08-480-640A-193
23	39.5	43.9	677	3	US-08-295-802-115
24	39.5	43.9	677	4	US-08-686-968C-58
25	39.5	43.9	677	4	US-08-686-968C-193
26	39.5	43.9	677	4	US-08-488-237A-115
27	39.5	43.9	677	4	US-08-488-237A-193

28	39.5	43.9	677	4	US-08-375-992A-115	Sequence 115, App
29	39.5	43.9	677	4	US-08-375-992A-193	Sequence 193, App
30	39	43.3	114	1	US-08-031-399-3	Sequence 3, App11
31	39	43.3	114	1	US-08-031-399-6	Sequence 6, App11
32	39	43.3	114	1	US-08-031-399-12	Sequence 12, App1
33	39	43.3	114	1	US-08-393-305-3	Sequence 3, App11
34	39	43.3	114	1	US-08-393-305-6	Sequence 6, App11
35	39	43.3	114	1	US-08-726-817-3	Sequence 3, App11
36	39	43.3	114	1	US-08-726-817-6	Sequence 6, App11
37	39	43.3	114	1	US-08-504-042-3	Sequence 3, App11
38	39	43.3	114	1	US-08-504-042-6	Sequence 6, App11
39	39	43.3	114	1	US-08-504-042-12	Sequence 12, App1
40	39	43.3	114	2	US-08-725-969-3	Sequence 3, App11
41	39	43.3	114	2	US-08-725-969-6	Sequence 6, App11
42	39	43.3	114	2	US-08-794-524-3	Sequence 3, App11
43	39	43.3	114	2	US-08-794-524-6	Sequence 6, App11
44	39	43.3	114	4	US-09-189-193-3	Sequence 3, App11
45	39	43.3	114	4	US-09-189-193-6	Sequence 6, App11

ALIGNMENTS

RESULT 1
US-07-725-331-62
Sequence 62, Application US/07725331

Patent No. 5294605

GENERAL INFORMATION:

APPLICANT: Houghten, Richard

TITLE OF INVENTION: Amphiphilic Peptide Compositions and

TITLE OF INVENTION: Analogues Thereof

NUMBER OF SEQUENCES: 68

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Dressler, Goldsmith, Sutter, Shore,

ADDRESSEE: 6 Milnamow

STREET: 180 No. 5294605th Stetson

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/725,331

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/554,422

FILING DATE: 19-JUL-1990

ATTORNEY/AGENT INFORMATION:

NAME: Gamson, Edward P.

REGISTRATION NUMBER: 29,381

REFERENCE/DOCKET NUMBER: 421250-80

TELECOMMUNICATION INFORMATION:

TELEPHONE: 3126165418

TELEFAX: 3126165460

INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids

TYPE: amino acids

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

OTHER INFORMATION: acetylated at N-terminus, may be

OTHER INFORMATION: a C-terminal amide.

US-07-725-331-62

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/05047
FILING DATE: 19910717
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
PCT-US91-05047-63

Query Match 45.6%; Score 41; DB 5; Length 36;
Best Local Similarity 31.2%; Pred. No. 5.1;
Matches 5; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

OY 1 KNLRRIRKIIHIKK 16
1 |:::|:::|
Db 9 KKLKKLKKLKKLKK 24

RESULT 5
US-09-156-316-4
; Sequence 4, Application US/09156316
; Patent No. 6183961
; GENERAL INFORMATION:
; APPLICANT: Bernstein, Harold S.
; APPLICANT: Coughlin, Shaun R.
; TITLE OF INVENTION: Methods and Compositions for Regulating Cell Cycle
; FILE REFERENCE: UCSF-020/0105
; CURRENT APPLICATION NUMBER: US/09/156,316
; CURRENT FILING DATE: 1998-09-18
; EARLIER APPLICATION NUMBER: 60/060,688
; EARLIER FILING DATE: 1997-09-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-156-316-4

Query Match 45.6%; Score 41; DB 4; Length 50;
Best Local Similarity 54.5%; Pred. No. 7;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 8 KRIIHIKKYG 18
:|:|:|:|:|:|
Db 11 QKVIELVKKYG 21

RESULT 6
US-07-820-154A-2
; Sequence 2, Application US/07820154A

Patent No. 5382425
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; APPLICANT: Junker M.S., David E
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,154A
; FILING DATE: 19920113
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-820-154A-2

Query Match 43.9%; Score 39.5; DB 1; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

OY 1 KNLRRIRKIIHI--IKK 16
||:|:|:|:|
Db 8 KNAKVIKSIISLQDIKK 26

RESULT 7
US-07-820-154A-6
; Sequence 6, Application US/07820154A
; Patent No. 5382425
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; APPLICANT: Junker M.S., David E
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,154A
; FILING DATE: 19920113
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P

APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,640A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-640A-2

Query Match 43.9%; Score 39.5; DB 3; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
QY 1 KNLRRIRKIITHI---IKK 16
|||:| | | | |
Db 8 KNARKVYSKIISLDIDIK 26

RESULT 11
US-08-480-640A-6
Sequence 6, Application US/08480640A
Patent No. 6033904
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
ADDRESSEE: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,640A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Swinepox virus
STRAIN: Kasza
POSITION IN GENOME:
MAP POSITION: ~23.2
UNITS: %G
US-08-480-640A-6

Query Match 43.9%; Score 39.5; DB 3; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
QY 1 KNLRRIRKIITHI---IKK 16
|||:| | | | |
Db 8 KNARKVYSKIISLDIDIK 26

RESULT 12
US-08-295-802-2
Sequence 2, Application US/08295802
Patent No. 6127163
GENERAL INFORMATION:
APPLICANT: Cochran Ph.D., Mark D
ADDRESSEE: Junker M.S., David E
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 188
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/295,802
FILING DATE: Herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-295-802-2

Query Match 43.9%; Score 39.5; DB 3; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

OY 1 KNLRIIRKIIH---IKK 16
|||:| ||| : |||
Db 8 KNAKVISKIISLQLDIKK 26

RESULT 13

US-08-295-802-6
; Sequence 6, Application US/08295802
; Patent No. 6127163
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 188
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/295,802
; FILING DATE: Herewith
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Swinepox virus
; STRAIN: Kasza
; POSITION IN GENOME:
; MAP POSITION: ~23.2
; UNITS: %G
US-08-295-802-6

Query Match 43.9%; Score 39.5; DB 3; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

OY 1 KNLRIIRKIIH---IKK 16
|||:| ||| : |||
Db 8 KNAKVISKIISLQLDIKK 26

RESULT 14
US-08-686-968C-102
; Sequence 102, Application US/08686968C
; Patent No. 6221361
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Junker, David E.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; FILE REFERENCE: 39119-H/JML

; CURRENT APPLICATION NUMBER: US/08/686,968C
; CURRENT FILING DATE: 1996-07-25
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 102
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Swinepox virus
US-08-686-968C-102

Query Match 43.9%; Score 39.5; DB 4; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

OY 1 KNLRIIRKIIH---IKK 16
|||:| ||| : |||
Db 8 KNAKVISKIISLQLDIKK 26

RESULT 15

US-08-488-237A-2
; Sequence 2, Application US/08488237A
; Patent No. 6251403
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Junker, David E.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,237A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-237A-2

Query Match 43.9%; Score 39.5; DB 4; Length 132;
Best Local Similarity 52.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

OY 1 KNLRIIRKIIH---IKK 16
|||:| ||| : |||
Db 8 KNAKVISKIISLQLDIKK 26

Search completed: July 12, 2002, 07:55:05
Job time: 112 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:56:49 ; Search time 95.47 Seconds
(without alignments)
18.117 Million cell updates/sec

Title: US-09-642-744B-7

Perfect score: 90
Sequence: 1 KNLRRIRKRIHIKKYG 18

Scoring table: BLUSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR-71:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	63.3	152	2	S68411 cathelin-related p
2	57	63.3	160	2	S68228 myeloid antimicrob
3	57	63.3	160	2	S68412 cathelin-related p
4	46	51.1	319	1	S19248 RNA-directed DNA p
5	45.5	50.6	381	1	G64416 conserved hypotbet
6	45	50.0	1634	2	E64410 DNA-directed DNA p
7	44	48.9	1670	2	T06754 DNA-directed RNA p
8	43.5	48.3	208	2	E81935 probable phosphori
9	43.5	48.3	208	2	G81169 N-(5'-phosphoribos
10	43	47.8	210	2	AC1577 precorrin isomerases
11	43	47.8	234	2	I40822 sigk protein - Cio
12	43	47.8	234	2	D97108 DNA-dependent RNA
13	43	47.8	234	2	D97108 DNA primase chain
14	43	47.8	420	2	S45630 conserved hypotbet
15	42	46.7	537	2	C90389 conserved hypotbet
16	42	46.7	41	2	S77768 hypotbetical prote
17	42	46.7	96	2	G64354 hypotbetical prote
18	42	46.7	125	2	D90581 hypotbetical prote
19	42	46.7	146	2	A72385 conserved hypotbet
20	42	46.7	152	2	A72385 arginine repressor
21	42	46.7	168	2	T08831 disease resistance
22	42	46.7	186	2	B90116 hypotbetical prote
23	42	46.7	187	2	H71819 hypotbetical prote
24	42	46.7	359	2	S77322 hypotbetical prote
25	42	46.7	429	2	D90428 hypotbetical prote
26	42	46.7	483	2	S75369 hypotbetical prote
27	42	46.7	507	2	T10753 intestinal DNA rep
28	42	46.7	540	2	T33962 hypotbetical prote
29	42	46.7	594	2	D70127 alanine--trna liga
30	42	46.7	698	2	D64084 helicase (EC 3.6.1

30	42	46.7	937	2	T37241 olfactory channel
31	42	46.7	957	2	D88651 protein B0212.5 (i
32	42	46.7	1741	2	T15978 hypotbetical prote
33	41.5	46.1	313	2	A89871 cmp-binding factor
34	41	45.6	117	2	T12722 hypotbetical prote
35	41	45.6	181	2	B90246 TATA box binding p
36	41	45.6	190	2	A85360 hypotbetical prote
37	41	45.6	198	2	S55311 TATA-binding prote
38	41	45.6	207	2	E70407 probable CDP-alcohol
39	41	45.6	256	2	A35340 H+-transporting AT
40	41	45.6	297	2	H70446 UTP--glucose-1-pho
41	41	45.6	337	2	S46010 hypotbetical prote
42	41	45.6	338	2	D90017 hypotbetical prote
43	41	45.6	367	2	S19172 cytochrome P450 2B
44	41	45.6	414	2	G64091 cell division prot
45	41	45.6	417	2	A33269 DNA primase (EC 2.

ALIGNMENTS

```
RESULT 1
S68411
cathelin-related protein 2 precursor - sheep (fragment)
C:Species: Ovis orientalis aries; Ovis ammon aries (domestic sheep)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68411
R:Mahoney, M.M.; Lee, A.Y.; Brezinski-Calguri, D.J.; Huttner, K.M.
FEBS Lett. 377, 519-522, 1995
A>Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobia
A:Reference number: S68411; MUID:96140561
A:Accession: S68411
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-152 <MAH>
A:Cross-references: EMBL:X92757
C:Genetics:
A:Gene: SC5-2
C:Superfamily: cathelin; cystatin homology
F:1-21/Domain: signal sequence #status predicted <SIG>
F:14-122/Domain: cystatin homology <CYS>
F:21-123/Domain: propeptide #status predicted <PRO>
F:124-152/Product: cathelin-related protein 2 #status predicted <MAT>

Query Match      63.3%; Score 57; DB 2; Length 152;
Best Local Similarity 61.1%; Pred. No. 0.19;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KNLRRIRKRIHIKKYG 18
Db 124 KGLRRIRKRIHIKKYG 141

RESULT 2
S68228
myeloid antimicrobial peptide 29 precursor - sheep
N:Alternate names: cathelicidin
C:Species: Ovis orientalis aries; Ovis ammon aries (domestic sheep)
C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
A:Accession: S68228
R:Bagella, L.; Scofield, M.; Zanetti, M.
FEBS Lett. 376, 225-228, 1995
A>Title: cDNA sequences of three sheep myeloid cathelicidins.
A:Reference number: S68228; MUID:96105386
A:Accession: S68228
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-160 <BAG>
A:Cross-references: EMBL:L46854; MUID:g1161248; PTDN:AAA65470.1; PTD:g1161249
C:Superfamily: cathelin; cystatin homology
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-130/Domain: cystatin homology <CYS>
```

F:29-131/Domain: propeptide #status predicted <PRO>
F:132-160/Product: myeloid antimicrobial peptide 29 #status predicted <MAT>

Query Match 63.3%; Score 57; DB 2; Length 160;
Best Local Similarity 61.1%; Pred. No. 0.2;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNLRRIRKIIHIIRKYG 18
: |||: |||: |||:
Db 132 RGLRRLGKRIAGVKKYG 149

RESULT 3
S68412
cathelin-related protein 1 precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68412
R:Mahoney, M.M.; Lee, A.Y.; Brezinski-Caliguri, D.J.; Huttner, K.M.
FEBS Lett. 377, 519-522, 1995
A:Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobial F
A:Reference number: S68411; MUID:96140581
A:Accession: S68412
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-160 <MAH>
A:Cross-references: EMBL:X92758
C:Genetics:
A:Gene: SC5-1
C:Superfamily: cathelin; cystatin homology
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-130/Domain: cystatin homology <CYS>
F:30-131/Domain: propeptide #status predicted <PRO>
F:132-160/Product: cathelin-related protein 1 #status predicted <MAT>

Query Match 63.3%; Score 57; DB 2; Length 160;
Best Local Similarity 61.1%; Pred. No. 0.2;
Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNLRRIRKIIHIIRKYG 18
: |||: |||: |||:
Db 132 RGLRRLGKRIAGVKKYG 149

RESULT 4
S19248
RNA-directed DNA polymerase (EC 2.7.7.49), msDNA specific - Escherichia coli retron Ec10
N:Alternate names: DNA nucleotidyltransferase (RNA-directed); reverse transcriptase; rev
C:Species: Escherichia coli retron Ec107
C:Date: 20-Feb-1995 #sequence_revision 15-Oct-1996 #text_change 11-Jun-1999
C:Accession: S19248
R:Herzer, P.J.; Inouye, S.; Inouye, M.
Mol. Microbiol. 6, 345-354, 1992
A:Title: Retron Ec107 is inserted into the Escherichia coli genome by replacing a palind
A:Reference number: S19248; MUID:92204001
A:Accession: S19248
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <HER>
A:Cross-references: EMBL:X62583; NID:q42774; PIDN:CA444468.1; PID:q42775
A:Experiment source: E. coli wild strain
C:Genetics:
A:Note: Insertion site is 82 min of E. coli K12 genetic map
C:Superfamily: reverse transcriptase
C:Keywords: nucleotidyltransferase

Query Match 51.1%; Score 46; DB 1; Length 319;
Best Local Similarity 38.9%; Pred. No. 18;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 KNLRRIRKIIHIIRKYG 18
: ||: ||: ||: |||:
Db 243 REARRALROEVHLCQKYG 260

RESULT 5
C64416
conserved hypothetical MG372 related protein - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: C64416
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A:Reference number: A64300; MUID:96337999
A:Accession: C64416
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-381 <BU>
A:Cross-references: GB:U67536; GB:L77117; NID:q1591596; PIDN:AAB98933.1; PID:q1591602
C:Genetics:
A:Map position: FOR860923-862068
C:Superfamily: Mycoplasma genitalium hypothetical protein MG372

Query Match 50.6%; Score 45.5; DB 1; Length 381;
Best Local Similarity 55.6%; Pred. No. 25;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 1 KNLRRIRKIIHIIRKYG 17
: ||: ||: ||: ||: ||:
Db 27 KNLRRIRKIIHIIRKYG 44

RESULT 6
E64410
DNA-directed DNA polymerase (EC 2.7.7.7) family B, intein containing precursor - Meth
N:Contains: DNA endonuclease (EC 3.1.1.-) PI-I; DNA endonuclease (EC 3.1.1.-) PI-II;
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Sep-1998
C:Accession: E64410
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A:Reference number: A64300; MUID:96337999
A:Accession: E64410
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1634 <BU>
A:Cross-references: GB:U67532; GB:L77117; NID:q1591599; PID:q1591563; TIGR:MJ0885; PI
A:Genetics:
A:Map position: REV816304-811400
A:Start codon: TTG
C:Function: <DPL>
A:Description: as DNA-directed DNA polymerase, catalyzes the polymerization of DNA at
C:Function: <EN1>
A:Description: as DNA endonuclease PI-Maj, catalyzes the hydrolysis of internal phos
C:Function: <EN2>
A:Description: as DNA endonuclease PI-Maj, catalyzes the hydrolysis of internal pho
A:Superfamily: as DNA-directed DNA polymerase, KOD
C:Keywords: endonuclease; hydrolase; nucleotidyltransferase; protein splicing
F:1-425/Domain: DNA-directed DNA polymerase family B extein 1 #status predict
F:1-425/Domain: DNA-directed DNA polymerase family B extein 1 #status predict <XT1>
F:426-794/Product: DNA endonuclease PI-I (pol B extein 1) #status predicted <MAT2>
F:795-882/Domain: DNA-directed DNA polymerase family B extein 2 #status predicted <XT
F:883-1358/Product: DNA endonuclease PI-II (pol B extein 2) #status predicted <MAT3>
F:1358-1634/Domain: DNA-directed DNA polymerase family B extein 3 #status predicted <

F:425-795/Cross-link: peptide (Arg-Ser) #status predicted
F:882-1359/Cross-link: peptide (Asn-Ser) #status predicted

Query Match 50.0%; Score 45; DB 2; Length 1634;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNLRIIRIKIHIKKY 17
||:|||||:|:|
DB 207 KNEKELIKRIETLKEY 223

RESULT 7

DNA-directed RNA polymerase I 190K chain homolog F15B8.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Mar-2000
C:Accession: T06754
R:Quetier, F.; Benes, V.; Rechmann, S.; Borckova, D.; Ansgorge, W.; Salanoubat, M.; Mewes, submitted to the Protein Sequence Database, April 1999
A:Reference number: 215794
A:Accession: T06754
A:Molecule type: DNA
A:Residues: 1-1670 <QUE>
A:Cross-references: EMBL:AL049660; GSPDB:GN00061; ATSP:F15B8.150
C:Experimental source: cultivar Columbia; BAC clone F15B8
C:Genetics:
A:Gene: ATSP:F15B8.150
A:Map position: 3
A:Introns: 21/3; 78/2; 131/3; 233/3; 306/3; 362/3; 417/1; 479/3; 567/3; 597/2; 646/3; 71
C:Superfamily: human DNA-directed RNA polymerase II largest chain

Query Match 48.9%; Score 44; DB 2; Length 1670;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 IIRKIIHIKKYG 18
|||||:|:|
DB 1569 IIREINHVKSYG 1581

RESULT 8

Probable phosphoribosylanthranilate isomerase (EC 5.3.1.24) NMA0890 [imported] - Neisseria
E81935
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: E81935
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holtroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis z2491.
A:Reference number: AB1775; MUID:20222556
A:Accession: E81935
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-208 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84170.1; PID:g737960
A:Experimental source: serogroup A, strain z2491
C:Genetics:
A:Gene: trpF; NMA0890
C:Superfamily: phosphoribosylanthranilate isomerase; trpF homology
C:Keywords: Intramolecular oxidoreductase; isomerase

Query Match 48.3%; Score 43.5; DB 2; Length 208;
Best Local Similarity 47.4%; Pred. No. 28;
Matches 9; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 1 KNLRIIRIKI-IHIKKYG 18
||:||||:|:|
DB 68 QNIRRIIAEVPRIHIQPHG 86

RESULT 9

N-(5'-phosphoribosyl)anthranilate isomerase NMB0688 [imported] - Neisseria meningitidis
G81169
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: G81169
R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizzato, M. Science 287, 1809-1815, 2000
A:Authors: Grandt, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappaport, R.; A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:2015755
A:Accession: G81169
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-208 <NET>
A:Cross-references: GB:AE002423; GB:AE002098; NID:g7225913; PIDN:AAF41106.1; PID:g7722
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0688
C:Superfamily: phosphoribosylanthranilate isomerase; trpF homology

Query Match 48.3%; Score 43.5; DB 2; Length 208;
Best Local Similarity 47.4%; Pred. No. 28;
Matches 9; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 1 KNLRIIRIKI-IHIKKYG 18
||:||||:|:|
DB 68 QNIRRIIAEVPRIHIQPHG 86

RESULT 10

precorrin isomerase homolog lin156 [imported] - Listeria innocua (strain Cljpl1262)
AC1577
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AC1577
R:Glaser, P.; Frangeul, L.; Buchliesser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schlueiter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1577
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96387.1; PID:g1641615; GSPDB:GN00178
A:Experimental source: strain Cljpl1262
C:Genetics:
A:Gene: lin156
C:Superfamily: Methanobacterium thermoautotrophicum precorrin isomerase

Query Match 47.8%; Score 43; DB 2; Length 210;
Best Local Similarity 53.8%; Pred. No. 33;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 6 IIRKIIHIKKYG 18
||:|||||:|
DB 62 VIOKIIHLKKNKG 74

RESULT 11

sigK protein - Clostridium acetobutylicum
140822
C:Species: Clostridium acetobutylicum

C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 21-Jul-2000
C/Accession: I40822
R:Sauer, U.; Treuner, A.; Buchholz, M.; Santangelo, J.D.; Durre, P.
J. Bacteriol. 176, 6572-6582, 1994
A>Title: Sporulation and primary sigma factor homologous genes in *Clostridium acetobutylicum*
A:Reference number: I40609; MUID:95050216
A:Accession: I40822
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-234 <RES>
A:Cross-references: GB:I23317; NID:g2745925; PIDN:AA894775.1; PID:g528974
C:Genetics:
A:Gene: sigK
A:Start codon: GTG
C:Superfamily: transcription initiation factor sigma K; transcription initiation factor
F:56-230/Domain: transcription initiation factor sigma Kalf homology <KTF>

Query Match 47.8%; Score 43; DB 2; Length 234;
Best Local Similarity 46.7%; Pred. No. 37;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 3 LRRIRKIIHIKKY 17
: | : |||
Db 55 IERNLRVLAHIKKY 69

RESULT 12
D97108
DNA-dependent RNA polymerase sigma chain [imported] - *Clostridium acetobutylicum*
C:Species: *Clostridium acetobutylicum*
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C/Accession: D97108
R:Molling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum*
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D97108
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK79655.1; PID:g15024653; GSPDB:GN00168
C:Genetics:
A:Experimental source: *Clostridium acetobutylicum* ATCC824
A:Gene: CAC1689
C:Superfamily: transcription initiation factor sigma K; transcription initiation factor

Query Match 47.8%; Score 43; DB 2; Length 234;
Best Local Similarity 46.7%; Pred. No. 37;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 3 LRRIRKIIHIKKY 17
: | : |||
Db 55 IERNLRVLAHIKKY 69

RESULT 13
S45630
DNA primase chain p48 - human
C:Species: *Homo sapiens* (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C/Accession: S45630
R:Stadlbauer, F.; Brueckner, A.; Rehtness, C.; Eckerskorn, C.; Lottspeich, F.; Foerster, Eur. J. Biochem. 222, 781-793, 1994
A>Title: DNA replication in vitro by recombinant DNA-polymerase-alpha-primase.
A:Reference number: S45628; MUID:94298818
A:Accession: S45630
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-420 <STA>
A:Cross-references: EMBL:X74330; NID:g510405; PIDN:CAA52377.1; PID:g510406

C:Superfamily: DNA primase 50K chain

Query Match 47.8%; Score 43; DB 2; Length 420;
Best Local Similarity 52.9%; Pred. No. 64;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 KNLRIIRKIIHIKKY 17
: | : |||
Db 206 EKHPFKRSINIHIKKY 222

RESULT 14
C90389
conserved hypothetical protein [imported] - *Sulfolobus solfataricus*
C:Species: *Sulfolobus solfataricus*
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C/Accession: C90389
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch Jung, I.; Jeffries, A.C.; Kozera, C.O.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder arett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
Submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A99139
A:Accession: C90389
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-537 <KUR>
A:Cross-references: GB:AE006641; NID:g13815498; PIDN:AAK42370.1; GSPDB:GN00155
A:Gene: SSO2200

Query Match 47.8%; Score 43; DB 2; Length 537;
Best Local Similarity 43.8%; Pred. No. 81;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 NLRRIIRKIIHIKKY 17
: || : |||
Db 81 DMNTIKATLHIKKY 96

RESULT 15
S77768
hypothetical protein MC037 - *Mycoplasma capricolum* (fragment)
C:Species: *Mycoplasma capricolum*
C>Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 21-Jul-2000
C/Accession: S77768; S48590
R:Bork, P.; Ouzounis, C.; Casati, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, M.O. Microbiol. 16, 955-967, 1995
A>Title: Exploring the *Mycoplasma capricolum* genome: a minimal cell reveals its physiology
A:Reference number: S77739; MUID:96059641
A:Accession: S77768
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-41 <BOR>
A:Cross-references: EMBL:Z33033; NID:g541693; PIDN:CAA83711.1; PID:g4379130
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Genetic code: SGC3

Query Match 46.7%; Score 42; DB 2; Length 41;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 KNLRIIRKIIHIKKY 16
: | : |||
Db 17 OELKEFIQWIIHIKK 32

Fri Jul 12 08:56:25 2002

us-09-642-744b-7.rpr

Page 5

Search completed: July 12, 2002, 07:56:49
Job time: 216 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:04:36 : Search time 46.15 Seconds

(without alignments)
15.102 Million cell updates/sec

Title: US-09-642-744B-7

Perfect score: 90

Sequence: 1 KNLRRIRKRIHIIRKYG 18

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	63.3	160	1	SC51_SHEEP
2	57	63.3	160	1	SC52_SHEEP
3	45.5	50.6	381	1	THIT_METJA
4	45	50.0	1634	1	DPOL_METJA
5	43	47.8	420	1	PRIL_HUMAN
6	43	47.8	821	1	MCMB_HUMAN
7	42	46.7	93	1	Y435_METJA
8	42	46.7	152	1	ARCR_THEMA
9	42	46.7	507	1	MCMB_RAT
10	42	46.7	594	1	SYA_BORBU
11	42	46.7	670	1	REP_HAEIN
12	41	45.6	138	1	R81_ARATH
13	41	45.6	159	1	MB28_BOVIN
14	41	45.6	198	1	TBP_SULSH
15	41	45.6	198	1	TBP_SULSO
16	41	45.6	256	1	ATPF_RAT
17	41	45.6	337	1	YBZ1_YEAST
18	41	45.6	414	1	FTSY_HAEIN
19	41	45.6	417	1	PRIL_MOUSE
20	41	45.6	491	1	CPB1_RAT
21	41	45.6	491	1	CPB2_RAT
22	41	45.6	613	1	DEAD_HAEIN
23	41	45.6	686	1	MYBB_CHICK
24	41	45.6	700	1	MYBB_HUMAN
25	41	45.6	704	1	MYBB_MOUSE
26	40	44.4	105	1	Y795_PYRO
27	40	44.4	189	1	TBP_THECE
28	40	44.4	197	1	TBP_SULAC
29	40	44.4	492	1	CPBC_RAT
30	40	44.4	552	1	YBUD_ECOLI
31	40	44.4	743	1	MYBB_XENLA
32	40	44.4	821	1	PPSA_PYRHO
33	39.5	43.9	584	1	LMRA_LACLA

34	39.5	43.9	584	1	LMRA_LACLA	P97046 lactococcus
35	39	43.3	108	1	YF24_METJA	O58919 methanococc
36	39	43.3	156	1	Y330_RICPR	O92498 rickettsia
37	39	43.3	162	1	IL15_CERAE	P40221 ceratophthec
38	39	43.3	162	1	IL15_FELCA	O97687 felis silve
39	39	43.3	162	1	IL15_HUMAN	P40933 homo sapien
40	39	43.3	162	1	IL15_MACMU	P48092 macaca mula
41	39	43.3	162	1	IL15_PIG	O95253 sus scrofa
42	39	43.3	190	1	TBP_PYRKO	O52366 pyrococcus
43	39	43.3	304	1	TF2A_YEAST	P20459 saccharomyc
44	39	43.3	320	1	RT86_ECOLI	P23070 escherichia
45	39	43.3	358	1	Y074_METJA	O60380 methanococc

ALIGNMENTS

RESULT ID	SC51_SHEEP	STANDARD	PRT	160 AA
AC	P49928:			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Cathelin-related peptide SC5 precursor 1 (Antibacterial peptide SMAP-29).			
DE	29) (Myeloid antibacterial peptide SMAP-29).			
OS	Ovis aries (Sheep)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Ovis.			
OX	NCBI_Taxid:9940;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE-Bone marrow;			
RA	MEDLINE-96140581; PubMed-8549789;			
RA	Manoney M.M., Lee A.Y., Brezinski-Caliguri D.J., Hutner K.M.;			
RT	"Molecular analysis of the sheep cathelin family reveals a novel			
RT	antimicrobial peptide.";			
RL	FERS Lett. 377:519-522(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE-Liver;			
RA	Hutner K.M., Manoney M.M.;			
RL	Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.			
CC	-1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.			
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X97757; CAA63412.1; -			
DR	EMBL: U60600; AAB49715.1; -			
DR	Interpro: IPR001894; Cathelicidin.			
DR	Pfam: PF00666; Cathelicidins; 1.			
DR	Prodom: PD001838; Cathelicidin; 1.			
DR	PROSITE: PS00946; CATHELICIDINS_1; 1.			
DR	PROSITE: PS00947; CATHELICIDINS_2; 1.			
KW	Antibiotic; Signal.			
FT	SIGNAL 1 29			POTENTIAL.
FT	PROPEP 30 131			BY SIMILARITY.
FT	CHAIN 132 160			CATHELIN-RELATED PEPTIDE SC5.
FT	MOD_RES 30 30			PYROLIDONE CARBOXYLIC ACID
FT				(BY SIMILARITY).
FT	DISULFID 86 97			BY SIMILARITY.
FT	DISULFID 108 125			BY SIMILARITY.
SO	SEQUENCE 160 AA; 17786 MW; BD9B3859C432C249 CRC64;			

Query Match 63.3%; Score 57; DB 1; Length 160;
 Best Local Similarity 61.1%; Pred. No. 0.041;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KNLRRIIRKIIHIIRKYG 18
 : |||: ||| I : ||| ||
 DB 132 RGLRIRGRKIAHGKRYG 149

RESULT 2

SC52_SHEEP STANDARD; PRT: 160 AA.
 ID SC52_SHEEP
 AC P49929;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Cathelin-related peptide SC5 precursor 2 (Antibacterial peptide SMAP-29) (Myeloid antibacterial peptide SMAP-29).
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Bone marrow;
 RA Mahoney M.M., Lee A.Y., Brezinski-Caliaguri D.J., Huttner K.M.;
 RT "Molecular analysis of the sheep cathelin family reveals a novel antimicrobial peptide."
 RT FEBS Lett. 377:519-522(1995).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Bone marrow;
 RA MEDLINE=96105386; PubMed=7498547;
 RA Bagella L., Scocchi M., Zanetti M.;
 RT "cDNA sequences of three sheep myeloid cathelicidins."
 RT FEBS Lett. 376:225-228(1995).
 RL -1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.
 CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X92758; CAA63413.1; -;
 DR EMBL: L46854; AAB85470.1; -;
 DR InterPro: IPR001894; Cathelicidin.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00666; Cathelicidins; 1.
 DR ProDom: PD001838; Cathelicidin; 1.
 DR SMART: SM00043; CT; 1.
 DR PROSITE: PS00946; CATHELICIDINS_1; 1.
 DR PROSITE: PS00947; CATHELICIDINS_2; 1.
 KW Antibiotic; Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT PROPEP 30 131 BY SIMILARITY.
 FT CHAIN 132 160 CATHELIN-RELATED PEPTIDE SC5.
 FT MOD_RES 30 30 PYROGLUTAMINE CARBOXYLIC ACID
 (BY SIMILARITY).
 FT DISULFID 86 97 BY SIMILARITY.
 FT DISULFID 108 125 BY SIMILARITY.
 FT CONFLICT 28 28 S->R (IN REF. 2).
 FT SEQUENCE 160 AA; 17742 MW; 4FB98A0935B51F CRC64;

Query Match 63.3%; Score 57; DB 1; Length 160;
 Best Local Similarity 61.1%; Pred. No. 0.041;
 Matches 11; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KNLRRIIRKIIHIIRKYG 18
 : |||: ||| I : ||| ||
 DB 132 RGLRIRGRKIAHGKRYG 149

RESULT 3

THII_METUA STANDARD; PRT: 381 AA.
 ID THII_METUA
 AC Q58341;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable thiamine biosynthesis protein thii.
 GN THII OR MJ0931.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 RA Uutterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii."
 RT Science 273:1058-1073(1996).
 RL -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE THIAZOLE MOIETY (BY SIMILARITY).
 CC -1- PATHWAY: THIAMINE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- SIMILARITY: BELONGS TO THE THII FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U67536; AAB98933.1; -;
 DR TIGR: MJ0931; -;
 DR InterPro: IPR004114; THUMP.
 DR InterPro: IPR003720; Thii.
 DR Pfam: PF02568; Thii; 1.
 DR Pfam: PF02926; THUMP; 1.
 KW Thiamine biosynthesis; Complete proteome.
 SQ SEQUENCE 381 AA; 43436 MW; 0A31P1069DA3357B CRC64;

Query Match 50.6%; Score 45.5; DB 1; Length 381;
 Best Local Similarity 55.6%; Pred. No. 6;
 Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

OY 1 KNLRRIIRK-IIHIIRKY 17
 : |||: ||| I : ||| ||
 DB 27 KNLEILRKNIKILRKY 44

RESULT 4
 DPOL_METUA STANDARD; PRT: 1634 AA.
 ID DPOL_METUA
 AC Q58295;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)


```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA replication licensing factor MCM6 (P105MCM).
GN MCM6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97431351; PubMed=9286856;
RA Tsunaga H., Tabata N., Hosoya S., Tamura K., Endo Y., Nojima H.;
RA "HSCM6: a new member of the human MCM/PI family encodes a protein
RT homologous to fission yeast Mif5."
RL Genes Cells 2:381-399(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98184833; PubMed=9516426;
RA Holtorf H.P., Baack M., Richter A., Ritzi M., Knippers R.;
RA "Human protein MCM6 on HeLa cell chromatin."
RL J. Biol. Chem. 273:7320-7325(1998).
RN [3]
RP SEQUENCE OF 640-821 FROM N.A.
RX MEDLINE=97131582; PubMed=8977093;
RA Harvey C.B., Wang Y., Darmoul D., Phillips A., Mantel N.,
RA Swallow D.M.;
RA "Characterisation of a human homologue of a yeast cell division cycle
RT gene, MCM6, located adjacent to the 5' end of the lactase gene on
RL chromosome 2q21."
FEBS Lett. 398:135-135(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF A SINGLE ROUND OF DNA
CC REPLICATION DURING S PHASE. BINDS TO CHROMATIN DURING G1 AND
CC DEACCH FROM IT DURING S PHASE AS IF IT LICENSES THE CHROMATIN TO
CC REPLICATE.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D84557; BAA12699.1; -
DR EMBL: U46838; AAC50766.1; -
DR EMBL: U67284; AAB48165.1; -
DR EMBL: U67281; AAB48165.1; JOINED.
DR EMBL: U67282; AAB48165.1; JOINED.
DR EMBL: U67283; AAB48165.1; JOINED.
DR MIM: 601806; -
DR InterPro: IPR001208; MCM.
DR Pfam: PF00493; MCM.1.
DR SMART: SM00350; MCM.1.
DR PROSITE: PS00847; MCM.1; 1.
DR PROSITE: PS50051; MCM.2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein;
KW DNA replication; ATP-binding; Cell cycle.
FT DOMAIN 346 553 MCM.
FT NP_BIND 396 403 ATP (POTENTIAL).
FT CONFLICT 377 387 PKTGGTSLR -> SKDNRRDLSS (IN REF. 2).
FT CONFLICT 495 495 A -> T (IN REF. 2).
FT CONFLICT 738 738 MISSING (IN REF. 3).
FT CONFLICT 790 790 L -> P (IN REF. 2).
SO SEQUENCE 821 AA; 92889 MW; F94968EB25A3E501 CRC64;

```

Query Match 47.8%; Score 43; DB 1; Length 821;
 Best Local Similarity 50.0%; Pred. No. 31;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

QY 2 NLRRIKRIHIHKY 17
ID 1 111 1:11 : 1
Db 768 NKRRIEKVIRHLTHY 783

RESULT 7
Y435_METUA
ID Y435_METUA STANDARD; PRT; 93 AA.
AC Q57877;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0435.
GN M0435.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MDL=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.D., Fuhrmann J.L., Nguyen D.,
RA Uiterback L., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: BELONGS TO THE M. JANNASCHII M0126 / M0128 / M0141 /
CC M0435 / M0604 / M1215 / M1217 / M1305 / M1379 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67495; AAB98423.1; -
DR TIGR: M0435; -
DR InterPro: IPR002934; NTP_transf.
DR Pfam: PF01909; NTP_transf.2; 1.
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 93 AA; 10661 MW; AA6FD2014B942A68 CRC64;

```

Query Match 46.7%; Score 42; DB 1; Length 93;
 Best Local Similarity 52.9%; Pred. No. 5.3;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

QY 2 NLRRIKRIHIHKY 18
ID 1 111 1:11 : 1
Db 2 NINEIKRIIPILKHG 18

RESULT 8
ARGR_THEME
ID ARGR_THEME STANDARD; PRT; 152 AA.
AC Q9WV19;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Arginine repressor.
GN ARGR OR TM0371.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;

```

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RA Dimova D., Sakanyan V.;
 RT "The autoregulated ArgR protein of the hyperthermophilic bacterium
 RT Thermotoga neapolitana: thermostability and DNA binding properties.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RA MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Uterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of *Thermotoga maritima*.";
 RL Nature 399:323-329(1999).
 CC -!- FUNCTION: REGULATES ARGININE BIOSYNTHESIS GENES (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (potential).
 CC -!- SIMILARITY: BELONGS TO THE ARG R FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ132286; CAB45262.1; -
 DR EMBL: AF001717; AAD35458.1; -
 DR HSSP: OJ1408; 1B4B.
 DR TIGR: TM0371; -
 DR InterPro: IPR001669; Arg_repress.
 DR Pfam: PF01316; Arg_repressor.1.
 DR Pfam: PF02863; Arg_repressor.C.1.
 DR Prodom: PD007402; Arg_repress.1.
 KW Transcription regulation; DNA-binding; Trans-acting factor; Repressor;
 KW Complete proteome.
 SQ SEQUENCE 152 AA; 17222 MW; BCD5384934D490E9 CRC64;
 OY 1 KNLRIIRKIIH-----IKKYG 18
 DB 6 KRRELIRKIIHEKKISNOPIVEELKKYG 35
 RESULT 9
 MCM6_RAT STANDARD; PRT: 507 AA.
 ID MCM6_RAT
 AC 062724;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA replication licensing factor MCM6 (Intestinal DNA replication
 DE protein) (Fragment).
 GN MCM6.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HOLTZMAN; TISSUE=Intestine;
 RX MEDLINE=96011641; PubMed=7590274;
 RA Sykes D.E., Weiser M.M.;

RT "Rat intestinal crypt-cell replication factor with homology to early
 RT S-phase proteins required for cell division.";
 RL Gene 163:243-247(1995).
 RN [2]
 RP REVISIONS.
 RC STRAIN=HOLTZMAN;
 RA Sykes D.E.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF A SINGLE ROUND OF DNA
 CC REPLICATION DURING S PHASE. BINDS TO CHROMATIN DURING G1 AND
 CC DETACH FROM IT DURING S PHASE AS IF IT LICENSES THE CHROMATIN TO
 CC REPLICATE.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U1565; AAC18424.1; -
 DR InterPro: IPR001208; MCM.
 DR Pfam: PF00493; MCM.1.
 DR SMART: SM00350; MCM.1.
 DR PROSITE: PS00847; MCM_1.1.
 DR PROSITE: PS50051; MCM_2.1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 KW DNA replication; ATP-binding; Cell cycle.
 FT NON_TER 1 1
 FT DOMAIN 32 239 MCM.
 FT NP_BIND 82 89 ATP (POTENTIAL).
 SQ SEQUENCE 507 AA; 57369 MW; 03CE569FEFCDC654 CRC64;
 OY 2 NLRIIRKIIHIIKKY 17
 DB 454 NKKRIIEKVVHRLTH 469
 RESULT 10
 SYA_BORBU STANDARD; PRT: 594 AA.
 ID SYA_BORBU
 AC 051238;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine-tRNA ligase) (Alans).
 GN ALAS OR BB0220.
 OS Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCBI_TaxID=139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35210 / B31;
 RX MEDLINE=98065943; PubMed=9403685;
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
 RA Lachluga R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwin M.,
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
 RA Uterback T., Matthey L., McDonald L., Artlich P., Bowman C.,
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
 RA Smith H.O., Venter J.C.;
 RT "Genomic sequence of a Lyme disease spirochete, *Borrelia
 RT burgdorferi*.";
 RL Nature 390:580-586(1997).

```
CC -1- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) -> AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001132; AAC66604.1; -.
DR TIGR: BB0220; -.
DR InterPro: IPR002106; AA_trna_ligase_11.
DR InterPro: IPR002318; trna-synt_2c; 2.
DR Pfam: PF01411; trna-synt_2c; 2.
DR PRINTS: PR00980; TRNASYNTHALA.
DR PROSITE: PS00179; AA_TRNA_LIGASE_II_1; FALSE_NEG.
DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; FALSE_NEG.
KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KM Complete proteome.
SQ SEQUENCE 594 AA; 67773 MW; 2337B306FD02349A CRC64;

Query Match 46.7%; Score 42; DB 1; Length 594;
Best Local Similarity 56.2%; Pred. No. 33;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 LRRIKRIIHIKKYG 18
Db 307 LRRLIRSRIRYAKKLG 322

RESULT 11
REP_HAEIN STANDARD: PRT; 670 AA.
AC P44804;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ATP-dependent DNA helicase rep (EC 3.6.1.-).
GN REP OR H10649.
OS Haemophilus influenzae.
CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
CC Haemophilus.
RT NCB1_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Flieschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT Influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: REP HELICASE IS A SINGLE-STRANDED DNA-DEPENDENT
CC APPASE INVOLVED IN DNA REPLICATION. IT CAN INITIATE UNWINDING
CC AT A NICK IN THE DNA. IT BINDS TO THE SINGLE-STRANDED DNA AND
CC ACTS IN A PROGRESSIVE FASHION ALONG THE DNA IN THE 3' TO 5'
CC DIRECTION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UNRD SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32748; AAC22309.1; -.
DR HSSP: P09980; IUNA.
DR TIGR: H10649; -.
DR InterPro: IPR000212; UvrD-helicase.
DR Pfam: PF00580; UvrD-helicase; 1.
KM Hydroxylase; Helicase; DNA replication; ATP-binding; DNA-binding;
KM Complete proteome.
FT NP_BIND 22 29 ATP (POTENTIAL).
SQ SEQUENCE 670 AA; 77745 MW; 802A131251ED4E24 CRC64;

Query Match 46.7%; Score 42; DB 1; Length 670;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 RRIIRKRIIHIKKYG 18
Db 30 RYIIRKRIIHLIERKCG 44

RESULT 12
RRII_ARATH STANDARD: PRT; 138 AA.
AC P56802;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chloroplast 30S ribosomal protein S11.
GN RPS11.
OS Arabidopsis thaliana (Mouse-ear cress).
CC Chloroplast.
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCB1_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA.
RX MEDLINE=20039611; PubMed=10574454;
RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.,
RT "Complete structure of the chloroplast genome of Arabidopsis
RT thaliana.";
RL DNA Res. 6:283-290(1999).
CC -1- SIMILARITY: BELONGS TO THE S11P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AP000423; BAA84418.1; -.
DR InterPro: IPR001971; Ribosomal_S11.
DR Pfam: PF00411; Ribosomal_S11; 1.
DR Prodom: PD001010; Ribosomal_S11; 1.
DR PROSITE: PS00054; RIBOSOMAL_S11; 1.
KM Ribosomal protein; Chloroplast.
SQ SEQUENCE 138 AA; 15023 MW; 7A97E2B96CFC7898 CRC64;

Query Match 45.6%; Score 41; DB 1; Length 138;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 KNLRIIRKRIIHI 13
```


Search completed: July 12, 2002, 08:04:36
Job time: 648 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:03:44 ; Search time 174.7 Seconds
(without alignments)
17.824 Million cell updates/sec

Title: US-09-642-744B-7
Perfect score: 90
Sequence: 1 KNLRIIRKIIHIKKYG 18

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_oranelle:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	60.0	146	2	09AIH3
2	50	55.6	255	10	09LJW2
3	48	53.3	500	3	09PER0
4	46	51.1	319	2	005804
5	46	51.1	954	5	09G015
6	46	51.1	969	5	09G022
7	44	48.9	756	10	09JXR9
8	44	48.9	840	12	09DHM4
9	44	48.9	1670	10	09SVY0
10	43.5	48.3	208	16	09K0C6
11	43.5	48.3	208	16	09JVD1
12	43	47.8	210	16	09ZCL5
13	43	47.8	234	16	059264
14	43	47.8	379	2	09L7Z6
15	43	47.8	537	17	09JWF6
16	43	47.8	4564	5	077075

17	42	46.7	41	2	048971	048971 mycoplasma
18	42	46.7	125	10	09AVZ1	09avz1 guillardia
19	42	46.7	146	16	098013	098q13 mycoplasma
20	42	46.7	147	17	0976M0	0976m0 sulfolobus
21	42	46.7	168	10	096433	096433 glycine max
22	42	46.7	186	10	09AVW6	09avw6 guillardia
23	42	46.7	187	16	09ZJH9	09zjh9 helicobacte
24	42	46.7	233	2	09L7X2	09l7x2 clostridium
25	42	46.7	234	5	0952W1	0952w1 caenornabai
26	42	46.7	275	10	09LRL7	09lrl7 arabidopsis
27	42	46.7	359	16	P73385	P73385 synechocyst
28	42	46.7	429	17	097V09	097vq9 sulfolobus
29	42	46.7	460	5	09VL03	09vl03 drosophila
30	42	46.7	483	17	P65941	P65941 sulfolobus
31	42	46.7	530	3	09C2E6	09c2e6 neurospora
32	42	46.7	540	5	09TX26	09tx26 caenornabai
33	42	46.7	937	5	017469	017469 caenornabai
34	41.5	46.1	313	16	053606	053606 staphylococ
35	41.5	46.1	357	10	09XE30	09xe30 oryza sativ
36	41	45.6	104	2	045020	045020 borrelia bu
37	41	45.6	106	2	044868	044868 borrelia bu
38	41	45.6	117	9	080196	080196 methanobact
39	41	45.6	190	10	09SUH0	09suh0 arabidopsis
40	41	45.6	207	16	067288	067288 aquifex aeo
41	41	45.6	217	1	0977R2	0977r2 uncultured
42	41	45.6	256	11	09C007	09c0q7 mus musculu
43	41	45.6	290	2	09S0H6	09s0h6 borrelia bu
44	41	45.6	297	16	067602	067602 aquifex aeo
45	41	45.6	309	2	044823	044823 borrelia bu

ALIGNMENTS

RESULT 1
09AIH3 PRELIMINARY; PRT; 146 AA.
ID 09AIH3;
AC 09AIH3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN L10.
GN RPL10.
OS Candidatus Carsonella ruddii.
OC Bacteria; Proteobacteria; gamma subdivision; Candidatus Carsonella.
OX NCBI_TaxID=114186;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21125546; PubMed=11222582;
RA Clark M.A., Baumann L., Thao M.L., Moran N.A., Baumann P.;
RT "Degenerative Minimalism in the Genome of a Psyllid Endosymbiont."
RL J. Bacteriol. 183:1853-1861(2001).
DR EMBL: AF274444; AAK17073.1; "-"
SQ SEQUENCE 146 AA; 17749 MW; 4EF8D7D33EB17864 CRC64;

Query Match 60.0%; Score 54; DB 2; Length 146;
Best Local Similarity 55.6%; Pred. No. 1.2;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
OY 1 KNLRIIRKIIHIKKYG 18
Db 124 KNLRIIRKIIHIKKYG 141
RESULT 2
09LJW2 PRELIMINARY; PRT; 255 AA.
ID 09LJW2;
AC 09LJW2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SIMILARITY TO SERINE/THREONINE KINASE.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA:
 RA kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.:
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA:
 RX MEDLINE=20363099; PubMed=10907853;
 RA Nakamura Y.,
 RT Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
 RT TAC and BAC clones.";
 RL DNA Res. 7:217-221(2000).
 DR EMBL: AP000388; BAB02946.1;
 DR InterPro: IPR002902; DUF26.
 DR Pfam: PF01657; DUF26; 2.
 KW Kinase.
 SO SEQUENCE 255 AA; 29435 MW; F75B3E52AEC18CE CRC64;

Query Match 55.6%; Score 50; DB 10; Length 255;
 Best Local Similarity 50.0%; Pred. No. 7.9;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 KNLRIIRKIIHIKKYG 18
 Db 48 KNLNVIIRNISHLHRYG 65

RESULT 3
 O9P6R0 PRELIMINARY; PRT; 500 AA.
 AC O9P6R0;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE TRANSCRIPTION FACTOR IIB 70 KDA SUBUNIT.
 GN SPBC13E7.10C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972H-;
 RA Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.,
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL354632; CAB8985.1;
 DR InterPro: IPR004366; Cyclin.
 DR InterPro: IPR000812; TFIIB.
 DR Pfam: PF00382; transcript1_fac2; 2.
 DR PRINTS: PR00685; TIFACTORIIB.
 DR SMART: SM00385; CYCLIN; 2.
 SO SEQUENCE 500 AA; 56749 MW; 6CF1DF9ADA2B1B2 CRC64;

Query Match 53.3%; Score 48; DB 3; Length 500;
 Best Local Similarity 42.9%; Pred. No. 30;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 2 NLRIIRKIIHIKK 15
 Db 245 NFRSRVREVVHVK 258

RESULT 4

005804
 ID 005804 PRELIMINARY; PRT; 319 AA.
 AC 005804;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE RNA-DIRECTED DNA POLYMERASE FROM RETRON EC107 (EC 2.7.7.49)
 DE (REVERSE TRANSCRIPTASE).
 GN RT.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ECOR 70;
 RX MEDLINE=92204001; PubMed=1372675;
 RA Heizer P.J., Inouye S., Inouye M.;
 RT "Retron-EC107 is inserted into the Escherichia coli genome by
 RT replacing a palindromic 34bp intergenic sequence.";
 RL Mol. Microbiol. 6:345-354(1992)
 CC -I- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE = N
 CC PYROPHOSPHATE + DNA(N) (EC 2.7.7.49).
 CC -I- MISCELLANEOUS: RETRONS MAY BE THE ANCESTORS OF RETROVIRUS.
 CC -I- SIMILARITY: TO OTHER REVERSE TRANSCRIPTASES IN BACTERIA AND
 CC RETROVIRUSES.
 DR EMBL: X62583; CA444468.1;
 DR InterPro: IPR000477; RVTse.
 DR Pfam: PF00078; rvt; 1.
 KW RNA-directed DNA polymerase; Nucleotidyltransferase;
 KW Transposable element.
 SO SEQUENCE 319 AA; 36363 MW; BA8E54DA76AC0FF CRC64;

Query Match 51.1%; Score 46; DB 2; Length 319;
 Best Local Similarity 38.9%; Pred. No. 39;
 Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 1 KNLRIIRKIIHIKKYG 18
 Db 243 REARRALRQEVHLCKYK 260

RESULT 5
 O9G0I5 PRELIMINARY; PRT; 954 AA.
 ID O9G0I5;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ALANYL-TRNA SYNTHETASE.
 GN ALAS.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20542069; PubMed=11078517;
 RX Bunjun S., Stathopoulos C., Graham D., Min B., Kitabatake M.,
 RA Wang A.L., Wang C.C., Vivas C.P., Weiss L.M., Soll D.;
 RT "A dual-specificity aminoacyl-tRNA synthetase in the deep-rooted
 RT eukaryote Giardia lamblia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12997-13002(2000).
 DR EMBL: AF245445; AAG23137.1;
 DR InterPro: IPR002318; tRNA-synt_2c.
 DR Pfam: PF01411; tRNA-synt_2c; 2.
 DR PRINTS: PR00980; TRNASYNTHALA.
 KW Aminoacyl-tRNA synthetase.
 SO SEQUENCE 954 AA; 106800 MW; 2F39A8AFBE2C45B8 CRC64;

Query Match 51.1%; Score 46; DB 5; Length 954;
 Best Local Similarity 41.2%; Pred. No. 11e+02;

Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 2 NLRRIIRKIIHIKKYG 18
|||::|||:
Db 368 NLRNVLRRVFFHLVSRG 384

RESULT 6
Q9U022 PRELIMINARY; PRT; 969 AA.
AC Q9U022;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALANYL-TRNA SYNTHETASE (FRAGMENT).
GN ALAS.
OS Giardia lamblia (Giardia intestinalis).
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
OX NCBI_TaxID=5741;
RN [1]
RP SEQUENCE FROM N.A.
RA Chihade J.W., Brown J.R., Schimmel P., Ribas de Pouplana L.;
RT "Detection of an Intermediate Stage of Mitochondria Genesis.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF188719; AAF05594.1; -;
DR InterPro: IPR002318; tRNA-synt_2c.
DR Pfam: PF01411; tRNA-synt_2c; 4.
DR PRINTS: PRO00980; TRNASYNTHALA.
DR Aminoacyl-tRNA synthetase.
FT NON_TER 969
SQ SEQUENCE 969 AA; 108438 MW; 1A0BCB1EF3780C80 CRC64;

Query Match 51.1%; Score 46; DB 5; Length 969;
Best Local Similarity 41.2%; Pred. No. 1.1e+02;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 2 NLRRIIRKIIHIKKYG 18
|||::|||:
Db 383 NLRNVLRRVFFHLVSRG 399

RESULT 7
Q93XR9 PRELIMINARY; PRT; 756 AA.
AC Q93XR9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BG55 PROTEIN.
GN BG55.
OS Bruguiera gymnorhiza.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Rhizophoraceae; Bruguiera.
OX NCBI_TaxID=39984;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Banzal T., Hershkovits G., Katcoff D.J., Hanagata N., Dubinsky Z.,
RA Karube I.;
RT "Identification of mRNA transcripts differentially expressed in
RT response to high salinity by means of differential display in the
RT mangrove plant Bruguiera gymnorhiza.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB061795; BAB55653.1; -;
SQ SEQUENCE 756 AA; 84482 MW; 1597F3D46C3E0809 CRC64;

Query Match 48.9%; Score 44; DB 10; Length 756;
Best Local Similarity 40.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 4 RRIIRKIIHIKKYG 18
::|||::|||:
Db 201 KRLIYLLHLKKYG 215

RESULT 8
Q9DHN4 PRELIMINARY; PRT; 840 AA.
AC Q9DHN4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 79R PROTEIN.
GN 79R.
OS Yaba-like disease virus (YLDV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OX Yatapoxvirus.
OX NCBI_TaxID=132475;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21176366; PubMed=11277691;
RA Lee H.J., Essani K., Smith G.L.;
RT "The genome sequence of Yaba-like disease virus, a yatapoxvirus.";
RL Virology 281:170-192(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Lee H.J.;
RL Thesis (2000), Sir William Dunn School of Pathology, University of.
DR EMBL: AJ293568; CAC21317.1; -;
SQ SEQUENCE 840 AA; 97985 MW; FAB9A0A5BE744491 CRC64;

Query Match 48.9%; Score 44; DB 12; Length 840;
Best Local Similarity 52.9%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 KNLRIIRKIIHIKKY 17
::|||::|||:
Db 334 ENLKEISEKIFVKKY 350

RESULT 9
Q9SVY0 PRELIMINARY; PRT; 1670 AA.
AC Q9SVY0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DNA-DIRECTED RNA POLYMERASE I 190K CHAIN-LIKE PROTEIN (EC 2.7.7.6).
GN F15B8.150.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Benes V., Rechmann S., Borkova D., Ansoorge W., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C., Queller F., Salanoubat M.;
RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL049660; CAB41189.1; -;
DR InterPro: IPR000722; RNA_POL_A.
DR InterPro: IPR002879; RNA_POL_A2.
DR Pfam: PF00623; RNA_POL_A; 1.

OY 6 IIRKIIHIIKKG 18
:|||||:1
Db 62 VIQIHLVKNKG 74

RESULT 13

OY 059264 PRELIMINARY; PRT; 234 AA.
AC 059264;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE RNA POLYMERASE SIGMA FACTOR (DNA-DEPENDENT RNA POLYMERASE SIGMA SUBUNIT).
DE SIGK OR CAC1689.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
NCBI_TaxID=1488;

SEQUENCE FROM N.A.
RX MEDLINE=95050216; PubMed=7961408;
RA Sauer U., Treuner A., Buchholz M., Santangelo J.D., Durre P.;
RT "Sporeulation and primary sigma factor homologous genes in Clostridium acetobutylicum";
RL J. Bacteriol. 176:6572-6582(1994).
[2]

SEQUENCE FROM N.A.
RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hilt J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum";
RL J. Bacteriol. 183:4823-4838(2001).
-I- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND THEN IS RELEASED.
-I- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.

CC EMBL: I23317; AAB94775.1; -;
CC EMBL: AE007679; AAK79655.1; -;
DR HSSP: P00579; ISIG.
DR InterPro: IPR000943; Sigma70.
DR Pfam: PF00140; sigma70.1.
DR PROSITE: PS00715; SIGMA70.1.
DR PROSITE: PS00716; SIGMA70.2; UNKNOWN.1.
KM DNA-binding; DNA-directed RNA polymerase; Sigma factor; Sporulation; Transcription regulation; Complete proteome.
SQ SEQUENCE 234 AA; 26586 MW; 091DDDD27A22BC1E4 CRC64;

Query Match 47.8%; Score 43; DB 16; Length 234;

Best Local Similarity 46.7%; Pred. No. 81;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 3 LRRIIRKIIHIIKKY 17
:|:|:|:|:|:|
Db 55 IERNLRVAHYIKKY 69

RESULT 14

OY 09L726 PRELIMINARY; PRT; 379 AA.
AC 09L726;

DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE REMRON REVERSE TRANSCRIPTASE.
KM RTT.
OS Saimonella enteritidis.
SQ Plasmid low molecular weight plasmid I.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Saimonella.
NCBI_TaxID=592;
RN [1]
RE SEQUENCE FROM N.A.
RX MEDLINE=21189311; PubMed=11292805;
RA Rychlik I., Sebkova A., Gregorova D., Karpiskova R.;
RT "Low-Molecular-Weight Plasmid of Saimonella enterica Serovar Enteritidis Codes for Retron Reverse Transcriptase and Influences Plaque Resistance";
RL J. Bacteriol. 183:2852-2858(2001).
DR EMBL: AF218051; AAF72414.1; -;
DR InterPro: IPR000123; RNA_DNAPolYms.
DR InterPro: IPR000477; RVTse.
DR Pfam: PF00078; rvt.1.
DR PRINTS: PR00866; RNADNAPOLMS.
KM Plasmid, RNA-directed DNA polymerase.
SQ SEQUENCE 379 AA; 43958 MW; 4A33CD21F35F9CDE CRC64;

Query Match 47.8%; Score 43; DB 2; Length 379;

Best Local Similarity 50.0%; Pred. No. 1,3e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 KNRIIRKIIHIIKKG 18
| | | | | | | | | |
Db 254 KENRRYRLVYICKKOG 271

RESULT 15

OY 097WL6 PRELIMINARY; PRT; 537 AA.
AC 097WL6;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, last annotation update)
DE HYPOTHETICAL PROTEIN SSO2200.
GN SSO2200.

OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
NCBI_TaxID=2287;
RN [1]
RE SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Adayar M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Thériault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: AE006826; AAK42370.1; -;
DR InterPro: IPR002789; DUF87.
DR Pfam: PF01935; DUF87.1.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 537 AA; 60140 MW; 752F08818FED5057 CRC64;

Query Match 47.8%; Score 43; DB 17; Length 537;

Best Local Similarity 43.8%; Pred. No. 1,8e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 2 NLRRIIRKIIHIIKKY 17
:|:|:|:|:|:|
Db 81 DMNTIIRKATLHIIKKY 96

Search completed: July 12, 2002, 08:03:44
Job time: 630 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:00:41 ; Search time 227.32 Seconds
(without alignments)
7.818 Million cell updates/sec

Title: US-09-642-744b-9
Perfect score: 79
Sequence: 1 NRRIRKRIHKKV 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	16	22	AA1980.DAT
2	79	100.0	18	22	AA1981.DAT
3	79	100.0	20	22	AA1982.DAT
4	77	97.5	16	22	AA1983.DAT
5	77	97.5	18	22	AA1984.DAT
6	77	97.5	18	22	AA1985.DAT
7	77	97.5	20	22	AA1986.DAT
8	77	97.5	20	22	AA1987.DAT
9	66	83.5	14	22	AA1988.DAT
10	64	81.0	14	22	AA1989.DAT
11	64	81.0	22	22	AA1990.DAT

12	62	78.5	13	22	AA1991.DAT	Ovine SMAP 29 cath
13	59	74.7	13	22	AA1992.DAT	Ovine SMAP 29 cath
14	57	72.2	12	22	AA1993.DAT	Ovine SMAP 29 cath
15	47	59.5	18	22	AA1994.DAT	Ovine SMAP 29 cath
16	47	59.5	28	22	AA1995.DAT	Ovine SMAP 29 cath
17	47	59.5	29	22	AA1996.DAT	Ovine SMAP 29 cath
18	44	55.7	160	22	AA1997.DAT	Ovine SMAP 29 cath
19	42	53.2	291	22	AA1998.DAT	Ovine SMAP 29 cath
20	42	53.2	1038	22	AA1999.DAT	Ovine SMAP 29 cath
21	41.5	52.5	167	22	AA2000.DAT	Ovine SMAP 29 cath
22	41	51.9	613	22	AA2001.DAT	Ovine SMAP 29 cath
23	40	50.6	18	22	AA2002.DAT	Ovine SMAP 29 cath
24	40	50.6	224	22	AA2003.DAT	Ovine SMAP 29 cath
25	40	50.6	1634	21	AA2004.DAT	Ovine SMAP 29 cath
26	40	50.6	1634	21	AA2005.DAT	Ovine SMAP 29 cath
27	39.5	50.0	1634	21	AA2006.DAT	Ovine SMAP 29 cath
28	39.5	50.0	208	21	AA2007.DAT	Ovine SMAP 29 cath
29	39.5	50.0	208	21	AA2008.DAT	Ovine SMAP 29 cath
30	39.5	50.0	312	22	AA2009.DAT	Ovine SMAP 29 cath
31	39.5	50.0	313	22	AA2010.DAT	Ovine SMAP 29 cath
32	38.5	50.0	313	22	AA2011.DAT	Ovine SMAP 29 cath
33	39.5	50.0	313	22	AA2012.DAT	Ovine SMAP 29 cath
34	39	49.4	235	22	AA2013.DAT	Ovine SMAP 29 cath
35	39	49.4	414	22	AA2014.DAT	Ovine SMAP 29 cath
36	38	48.1	17	17	AA2015.DAT	Ovine SMAP 29 cath
37	38	48.1	17	17	AA2016.DAT	Ovine SMAP 29 cath
38	38	48.1	76	21	AA2017.DAT	Ovine SMAP 29 cath
39	38	48.1	149	20	AA2018.DAT	Ovine SMAP 29 cath
40	38	48.1	180	22	AA2019.DAT	Ovine SMAP 29 cath
41	38	48.1	229	21	AA2020.DAT	Ovine SMAP 29 cath
42	38	48.1	299	19	AA2021.DAT	Ovine SMAP 29 cath
43	38	48.1	457	20	AA2022.DAT	Ovine SMAP 29 cath
44	38	48.1	457	21	AA2023.DAT	Ovine SMAP 29 cath
45	38	48.1	491	22	AA2024.DAT	Ovine SMAP 29 cath

ALIGNMENTS

RESULT 1
ID AAB70657 standard; peptide: 16 AA.
XX AAB70657:
DT 15-MAY-2001 (first entry)
XX
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:10.
XX
XX
KW Ovine: SMAP29: lupine: RCAP 18: cathelicidin: antimicrobial:
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
XX
XX WO200112668-A1.
XX
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
XX
XX (IOWA) UNIV IOWA RES FOUND.
XX (REGC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections -
PT

XX Claim 1; Page 103; 137pp; English.
PS
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
SQ Sequence 16 AA;
XX
XX
Query Match 100.0%; Score 79; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NRRIRKIIHIIRKY 16
Db 1 NRRIRKIIHIIRKY 16
XXXXXXXXXXXXXXXXXXXX
RESULT 2
AAB70654 ID AAB70654 standard; peptide; 18 AA.
XX
XX AAB70654;
AC
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:7.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KM *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX Ovis aries.
OS
XX
PN WO200112668-A1.
XX
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI; 2001-234911/24.
DR
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
SQ Sequence 16 AA;
XX
XX
Query Match 100.0%; Score 79; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NRRIRKIIHIIRKY 16
Db 1 NRRIRKIIHIIRKY 16
XXXXXXXXXXXXXXXXXXXX
RESULT 2
AAB70654 ID AAB70654 standard; peptide; 18 AA.
XX
XX AAB70654;
AC
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:7.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KM *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX Ovis aries.
OS
XX
PN WO200112668-A1.
XX
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI; 2001-234911/24.
DR
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
SQ Sequence 16 AA;
XX
XX
Query Match 100.0%; Score 79; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NRRIRKIIHIIRKY 16
Db 1 NRRIRKIIHIIRKY 16
XXXXXXXXXXXXXXXXXXXX

CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
SQ Sequence 18 AA;
XX
XX
Query Match 100.0%; Score 79; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NRRIRKIIHIIRKY 16
Db 2 NRRIRKIIHIIRKY 17
XXXXXXXXXXXXXXXXXXXX
RESULT 3
AAB70653 ID AAB70653 standard; peptide; 20 AA.
XX
XX AAB70653;
AC
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:6.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KM *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX Ovis aries.
OS
XX
PN WO200112668-A1.
XX
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI; 2001-234911/24.
DR
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
SQ Sequence 20 AA;
XX
XX
Query Match 100.0%; Score 79; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NRRIRKIIHIIRKY 16
Db 1 NRRIRKIIHIIRKY 16
XXXXXXXXXXXXXXXXXXXX

```
Db      2 nrrlrkrlhikky 17
RESULT  4
AAB70656
ID      AAB70656 standard; peptide; 16 AA.
XX
AC      AAB70656;
XX
DT      15-MAY-2001 (first entry)
XX
DE      Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:9.
XX
KW      Ovine: SMAP29; lupine: RCAP 18; cathelicidin: antimicrobial;
KW      bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW      proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW      Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS      Ovis aries.
XX
PN      MO200112668-A1.
XX
PD      22-FEB-2001.
XX
PF      18-AUG-2000; 2000MO-US22781.
XX
PR      18-AUG-1999; 99US-0149886.
XX
PA      (IOWA ) UNIV IOWA RES FOUND.
PA      (REGC ) UNIV CALIFORNIA.
PI      Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX      WPI; 2001-234911/24.
XX
PT      New antimicrobial peptides useful as antibiotics for inhibiting growth
PT      and proliferation of microbes, and for treating microbial infections
XX
PS      Claim 1; Page 103; 137pp; English.
XX
CC      AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC      AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC      SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC      derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC      antibiotic, antimicrobial and antiviral activities, and can be used as
CC      microbial growth and proliferation inhibitors and in gene therapy. (I)
CC      are useful for inhibiting microbial growth in an environment capable of
CC      sustaining such growth, for inhibiting microbial growth or strain in a
CC      host, and inhibiting the growth of drug-resistant microbial strains such
CC      as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC      Xanthomonas.
XX
SQ      Sequence 16 AA:
XX
Query Match      97.5%; Score 77; DB 22; Length 16;
Best Local Similarity 93.8%; Pred. No. 6; 7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY      1 NRRIRKRIHIIRKY 16
      1:|||||
Db      1 nrrlrkrlhikky 16
RESULT  5
AAB70648
ID      AAB70648 standard; peptide; 18 AA.
XX
AC      AAB70648;
XX
DT      15-MAY-2001 (first entry)
XX
DE      Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:1.
XX
```

```
XX
KW      Ovine: SMAP29; lupine: RCAP 18; cathelicidin: antimicrobial;
KW      bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW      proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW      Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS      Ovis aries.
XX
PN      MO200112668-A1.
XX
PD      22-FEB-2001.
XX
PF      18-AUG-2000; 2000MO-US22781.
XX
PR      18-AUG-1999; 99US-0149886.
XX
PA      (IOWA ) UNIV IOWA RES FOUND.
PA      (REGC ) UNIV CALIFORNIA.
PI      Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX      WPI; 2001-234911/24.
XX
PT      New antimicrobial peptides useful as antibiotics for inhibiting growth
PT      and proliferation of microbes, and for treating microbial infections
XX
PS      Claim 1; Page 103; 137pp; English.
XX
CC      AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC      AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC      SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC      derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC      antibiotic, antimicrobial and antiviral activities, and can be used as
CC      microbial growth and proliferation inhibitors and in gene therapy. (I)
CC      are useful for inhibiting microbial growth in an environment capable of
CC      sustaining such growth, for inhibiting microbial growth or strain in a
CC      host, and inhibiting the growth of drug-resistant microbial strains such
CC      as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC      Xanthomonas.
XX
SQ      Sequence 18 AA:
XX
Query Match      97.5%; Score 77; DB 22; Length 18;
Best Local Similarity 93.8%; Pred. No. 7; 5e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY      1 NRRIRKRIHIIRKY 16
      1:|||||
Db      2 nrrlrkrlhikky 17
RESULT  6
AAB70655
ID      AAB70655 standard; peptide; 18 AA.
XX
AC      AAB70655;
XX
DT      15-MAY-2001 (first entry)
XX
DE      Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:8.
XX
KW      Ovine: SMAP29; lupine: RCAP 18; cathelicidin: antimicrobial;
KW      bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW      proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW      Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS      Ovis aries.
XX
PN      MO200112668-A1.
XX
PD      22-FEB-2001.
XX
```


PF 18-AUG-2000; 2000WO-US22781.
 XX
 PR 18-AUG-1999; 99US-0149886.
 XX
 PA (IOWA) UNIV IOWA RES. FOUND.
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R:
 XX
 DR WPI: 2001-234911/24.
 XX
 PT
 PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 and proliferation of microbes, and for treating microbial infections -
 XX
 XX
 PS Claim 1; Page 103; 137pp; English.
 XX
 CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664 and AAB70674 and AAB70675 are derived from the ovine
 CC SMP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.
 XX
 CQ Sequence 18 AA;
 XX

PT and proliferation of microbes, and for treating microbial infections -
XX
PS
XX Claim 1; Page 103; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCPAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 29 AA;

Query Match 97.5%; Score 77; DB 22; Length 29;
Best Local Similarity 93.8%; Pred. No. 1.2e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

QY 1 NRRIRRIIRIHIKKKY 16
|:|||||
Db 2 nrrirrkiihiikky 17

RESULT 8

Query Match	Score	DB	Length
Best Local Similarity	93.8%		
Matches	15: Conservative	1: Mismatches	0: Gaps
OY	1	NIRRIIRKIIHIIKKY	16
DB	2	nlrriirklrhikky	17
RESULT	7		
AAB70675			
ID	AAB70675	standard; peptide; 29 AA.	
XX	AAB70675;		
AC			
XX	15-MAY-2001	(first entry)	
DT			
DE	Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:28.		
XX			
KW	Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial; bactericidal; antibiotic; antiviral; microbial growth inhibitor; proliferation inhibitor; gene therapy; Pseudomonas aeruginosa; Burkholderia cepacia; Alcaligenes; Xanthomonas.		
XX			
OS	Ovis aries.		
XX			
PN	WO200112668-A1.		
XX			
PD	22-FEB-2001.		
XX			
PF	18-AUG-2000; 2000MO-US22781.		
PR	18-AUG-1999; 99US-0149886.		
XX			
PA	(IOWA) UNIV IOWA RES FOUND.		
XX	(RECC) UNIV CALIFORNIA.		
PI	Tack BE, McCreay P, Welsh M, Travis SM, Lehrer R;		
XX			
DR	WPI: 2001-234911/24.		
PT	New antimicrobial peptides useful as antibiotics for inhibiting growth		

XX AAB70660;
AC AAB70660;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:13.
XX
XX Ovine: SMAP29; lupine: RCAP 18; cathelicidin: antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES. FOUND.
PA (RESC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
DR WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
XX AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antimicrobial, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (II)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a

CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.

XX Sequence 14 AA;

Query Match 83.5%; Score 66; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.00034;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IRRIRKIIHIKK 15
 Db 1 IRRIRKIIHIKK 14

RESULT 9
 ID AAB70661 standard; peptide; 14 AA.

XX AAB70661;

DT 15-MAY-2001 (first entry)

XX Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:14.

XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;

XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;

XX proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;

XX *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

XX Ovis aries.

XX WO200112668-A1.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-US22781.

XX 18-AUG-1999; 99US-0149886.

XX (IOWA) UNIV IOWA RES FOUND.

XX (REGC) UNIV CALIFORNIA.

XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX WPI, 2001-234911/24.

XX Claim 1; Page 103; 137pp; English.

XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which

XX AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine

XX SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are

XX derived from the lupine RCAP 18 cathelicidin family peptide. (I) have

XX antibiotic, antimicrobial and antiviral activities, and can be used as

XX microbial growth and proliferation inhibitors and in gene therapy. (I)

XX are useful for inhibiting microbial growth in an environment capable of

XX sustaining such growth, for inhibiting microbial growth or strain in a

XX host, and inhibiting the growth of drug-resistant microbial strains such

XX as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and

XX *Xanthomonas*.

Db 1 IRRIRKIIHIKK 14

RESULT 10

ID AAB70658 standard; peptide; 14 AA.

XX AAB70658;

DT 15-MAY-2001 (first entry)

XX Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:11.

XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;

XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;

XX proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;

XX *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

XX Ovis aries.

XX WO200112668-A1.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-US22781.

XX 18-AUG-1999; 99US-0149886.

XX (IOWA) UNIV IOWA RES FOUND.

XX (REGC) UNIV CALIFORNIA.

XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX WPI, 2001-234911/24.

XX Claim 1; Page 103; 137pp; English.

XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which

XX AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine

XX SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are

XX derived from the lupine RCAP 18 cathelicidin family peptide. (I) have

XX antibiotic, antimicrobial and antiviral activities, and can be used as

XX microbial growth and proliferation inhibitors and in gene therapy. (I)

XX are useful for inhibiting microbial growth in an environment capable of

XX sustaining such growth, for inhibiting microbial growth or strain in a

XX host, and inhibiting the growth of drug-resistant microbial strains such

XX as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and

XX *Xanthomonas*.

XX Sequence 14 AA;

Query Match 81.0%; Score 64; DB 22; Length 14;
 Best Local Similarity 92.9%; Pred. No. 0.00072;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 IRRIRKIIHIKK 15
 Db 1 IRRIRKIIHIKK 14

RESULT 11
 ID AAB70659 standard; peptide; 14 AA.

XX AAB70659;

DT 15-MAY-2001 (first entry)

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
PS Claim 1; Page 103; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 13 AA;

Query Match 74.7%; Score 59; DB 22; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.0042;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRIIRKIIHIK 14
Db 1 IRIIRKIIHIK 13

RESULT 14
AAB70664
ID AAB70664 standard; peptide; 12 AA.
XX
NC AAB70664;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:17.
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
DR WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
XX
PS Claim 1; Page 103; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 13 AA;

CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 12 AA;

Query Match 72.2%; Score 57; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0082;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RRIIRKIIHIK 14
Db 1 IRIIRKIIHIK 12

RESULT 15
AAB70652
ID AAB70652 standard; peptide; 18 AA.
XX
NC AAB70652;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:5.
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
OS Ovis aries.
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
DR WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
XX
PS Claim 10; Page 107; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 18 AA;

Query Match 59.5%; Score 47; DB 22; Length 18;
Best Local Similarity 60.0%; Pred. No. 0.48;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 IRRIRKIRIKKY 16
:ll :lll l :lll
Db 3 Irlgrkiahgvkky 17

Search completed: July 12, 2002, 08:00:41
Job time: 448 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:55:05 ; Search time 75.52 Seconds
(without alignments)
5.175 Million cell updates/sec

Title: US-09-642-744B-9

Perfect score: 79

Sequence: 1 NRRIRKIHITIKRY 16

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	48.1	300	2	US-08-897-340-32
2	38	48.1	300	4	US-09-252-329-32
3	38	48.1	457	3	US-09-416-213-2
4	37	46.8	677	3	US-08-480-640A-115
5	37	46.8	677	3	US-08-480-640A-193
6	37	46.8	677	3	US-08-295-802-115
7	37	46.8	677	4	US-08-686-968C-58
8	37	46.8	677	4	US-08-686-968C-193
9	37	46.8	677	4	US-08-488-237A-115
10	37	46.8	677	4	US-08-488-237A-193
11	37	46.8	677	4	US-08-375-992A-115
12	37	46.8	677	4	US-08-375-992A-193
13	36	45.6	27	2	US-08-505-486-65
14	36	45.6	27	3	US-08-801-028-65
15	36	45.6	27	3	US-09-340-154-65
16	36	45.6	27	5	PCT-US95-09338-65
17	36	45.6	27	5	PCT-US95-09339-65
18	36	45.6	114	1	US-08-031-399-3
19	36	45.6	114	1	US-08-031-399-6
20	36	45.6	114	1	US-08-031-399-12
21	36	45.6	114	1	US-08-393-305-3
22	36	45.6	114	1	US-08-393-305-6
23	36	45.6	114	1	US-08-726-817-3
24	36	45.6	114	1	US-08-726-817-6
25	36	45.6	114	1	US-08-504-042-3
26	36	45.6	114	1	US-08-504-042-6
27	36	45.6	114	1	US-08-504-042-12

28	36	45.6	114	2	US-08-725-969-3	Sequence 3, Appl1
29	36	45.6	114	2	US-08-725-969-6	Sequence 6, Appl1
30	36	45.6	114	2	US-08-794-552-3	Sequence 3, Appl1
31	36	45.6	114	2	US-08-794-552-6	Sequence 6, Appl1
32	36	45.6	114	4	US-09-189-193-3	Sequence 3, Appl1
33	36	45.6	114	4	US-09-189-193-6	Sequence 6, Appl1
34	36	45.6	114	5	PCT-US94-03793-3	Sequence 3, Appl1
35	36	45.6	114	5	PCT-US94-03793-6	Sequence 6, Appl1
36	36	45.6	114	5	PCT-US94-03793-12	Sequence 12, Appl1
37	36	45.6	122	1	US-08-300-902A-3	Sequence 3, Appl1
38	36	45.6	162	1	US-08-031-399-2	Sequence 2, Appl1
39	36	45.6	162	1	US-08-031-399-5	Sequence 5, Appl1
40	36	45.6	162	1	US-08-393-305-2	Sequence 2, Appl1
41	36	45.6	162	1	US-08-393-305-5	Sequence 5, Appl1
42	36	45.6	162	1	US-08-535-733-2	Sequence 2, Appl1
43	36	45.6	162	1	US-08-284-393B-9	Sequence 9, Appl1
44	36	45.6	162	1	US-08-726-817-2	Sequence 2, Appl1
45	36	45.6	162	1	US-08-726-817-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1
US-08-897-340-32
; Sequence 32, Application US/08897340
; Patent No. 5955306
; GENERAL INFORMATION:
; APPLICANT: Gimeno, Carlos J. and Errada, Patrick, R.
; TITLE OF INVENTION: Weight Control Pathway Genes and Uses
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/897,340
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/715,032
; FILING DATE: 17-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Silver1, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-005CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-897-340-32

Query Match 48.1%; Score 38; DB 2; Length 300;
Best Local Similarity 33.3%; Pred. No. 79;
Matches 5; Conservative 8; Mismatches 2; Indels 0; Gaps 0;
QY 2 IRRIRKIHITIKRY 16

Db 234 VRLEVKQIHHMLREY 248

RESULT 2
US-09-252-329-32
Sequence 32, Application US/09252329
Patent No. 6147192
GENERAL INFORMATION:
APPLICANT: Gimeno, Carlos J. and Errada, Patrick, R.
TITLE OF INVENTION: Weight Control Pathway Genes and Uses
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/252,329
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/897,340
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Silverl, Jean M.
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: MNI-005CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-09-252-329-32

Query Match 48.1%; Score 38; DB 4; Length 300;
Best Local Similarity 33.3%; Pred. No. 79;
Matches 5; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

OY 2 IRRIRKIHIRKY 16
Db 234 VRLEVKQIHHMLREY 248

RESULT 3
US-09-416-213-2
Sequence 2, Application US/09416213
Patent No. 6110690
GENERAL INFORMATION:
APPLICANT: Goeddel, David V.
APPLICANT: Jiaq, Yinying
TITLE OF INVENTION: Suppressors of Death Domains
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA

ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/416,213
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/035,676
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: 798-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 457 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-416-213-2

Query Match 48.1%; Score 38; DB 3; Length 457;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 6 IRKIHIRK 15
Db 380 IKIHIVLEK 389

RESULT 4
US-08-480-640A-115
Sequence 115, Application US/08480640A
Patent No. 6033904
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,640A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 677 amino acids
TYPE: amino acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Swinepox virus
STRAIN: Kasza
INDIVIDUAL ISOLATE: S-SPV-001
IMMEDIATE SOURCE:
CLONE: 515-85.1
POSITION IN GENOME:
MAP POSITION: ~23.2
UNITS: %G
US-08-480-640A-115

Query Match 46.8%; Score 37; DB 3; Length 677;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 RIIRKIHIIKKY 16
| | : | | : | | |
Db 200 RDINRIIEILKKY 212

RESULT 5
US-08-480-640A-193
Sequence 193, Application US/08480640A
Patent No. 6033904
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480.640A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 193:
SEQUENCE CHARACTERISTICS:
LENGTH: 677 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-640A-193

Query Match 46.8%; Score 37; DB 3; Length 677;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 RIIRKIHIIKKY 16
| | : | | : | | |
Db 200 RDINRIIEILKKY 212

RESULT 6
US-08-295-802-115
Sequence 115, Application US/08295802
Patent No. 6127163
GENERAL INFORMATION:
APPLICANT: Cochran Ph.D., Mark D
APPLICANT: Junker M.S., David E
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 188
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/295,802
FILING DATE: Herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)977-9550
TELEFAX: (212)664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 677 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Swinepox virus
STRAIN: Kasza
INDIVIDUAL ISOLATE: S-SPV-001
IMMEDIATE SOURCE:
CLONE: 515-85.1
POSITION IN GENOME:
MAP POSITION: ~23.2
UNITS: %G
US-08-295-802-115

Query Match 46.8%; Score 37; DB 3; Length 677;
Best Local Similarity 53.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 RIIRKIHIIKKY 16
| | : | | : | | |
Db 200 RDINRIIEILKKY 212

RESULT 7
US-08-686-968C-58
Sequence 58, Application US/08686968C
Patent No. 6221361
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
FILE REFERENCE: 39119-H/JML
CURRENT APPLICATION NUMBER: US/08/686,968C


```
; CURRENT FILING DATE: 1996-07-25
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Swinepox virus
US-08-686-968C-58

Query Match
Best Local Similarity 46.8%; Score 37; DB 4; Length 677;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 RIIRKIHIKKY 16
Db 200 RDINRIIELEKKY 212

RESULT 8
US-08-686-968C-193
; Sequence 193, Application US/08686968C
; Patent No. 6221361
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Junker, David E.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; FILE REFERENCE: 39119-H/JML
; CURRENT APPLICATION NUMBER: US/08/686,968C
; CURRENT FILING DATE: 1996-07-25
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 193
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Swinepox virus
US-08-686-968C-193

Query Match
Best Local Similarity 46.8%; Score 37; DB 4; Length 677;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 RIIRKIHIKKY 16
Db 200 RDINRIIELEKKY 212

RESULT 9
US-08-488-237A-115
; Sequence 115, Application US/08488237A
; Patent No. 6251403
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Junker, David E.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,237A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 677 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Swinepox virus
; STRAIN: Kasza
; INDIVIDUAL ISOLATE: S-SPV-001
; IMMEDIATE SOURCE:
; CLONE: 515-85.1
; POSITION IN GENOME:
; MAP POSITION: ~23.2
; UNITS: %G
US-08-488-237A-115

Query Match
Best Local Similarity 46.8%; Score 37; DB 4; Length 677;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 RIIRKIHIKKY 16
Db 200 RDINRIIELEKKY 212

RESULT 10
US-08-488-237A-193
; Sequence 193, Application US/08488237A
; Patent No. 6251403
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D.
; APPLICANT: Junker, David E.
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,237A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 193:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 677 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

US-08-488-237A-193

Query Match 46.8%; Score 37; DB 4; Length 677;

Best Local Similarity 53.8%; Pred. NO. 2.4e+02; Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 RIIRKIHIIRKY 16
| | : | | : | | |
DB 200 RDINRIIEILKKY 212

RESULT 11

US-08-375-992A-115

; Sequence 115, Application US/08375992A
; Patent No. 6328975

; GENERAL INFORMATION:

; APPLICANT: Cochran, Mark D.

; APPLICANT: Junker, David E.

; TITLE OF INVENTION: Recombinant Swinepox Virus

; NUMBER OF SEQUENCES: 220

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John P. White

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/375,992A

; FILING DATE: Herewith

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P

; REGISTRATION NUMBER: 28,678

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 115:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 677 amino acids

; TYPE: amino acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Swinepox virus

; STRAIN: Kasza

; INDIVIDUAL ISOLATE: S-SPV-001

; IMMEDIATE SOURCE:

; CLONE: 515-85.1

; POSITION IN GENOME:

; MAP POSITION: -23.2

; UNITS: %G

; US-08-375-992A-115

Query Match 46.8%; Score 37; DB 4; Length 677;

Best Local Similarity 53.8%; Pred. NO. 2.4e+02; Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 RIIRKIHIIRKY 16
| | : | | : | | |
DB 200 RDINRIIEILKKY 212

RESULT 12

US-08-375-992A-193

; Sequence 193, Application US/08375992A
; Patent No. 6328975

; GENERAL INFORMATION:

; APPLICANT: Cochran, Mark D.

; APPLICANT: Junker, David E.

; TITLE OF INVENTION: Recombinant Swinepox Virus

; NUMBER OF SEQUENCES: 220

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John P. White

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/375,992A

; FILING DATE: Herewith

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P

; REGISTRATION NUMBER: 28,678

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 193:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 677 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-375-992A-193

Query Match 46.8%; Score 37; DB 4; Length 677;

Best Local Similarity 53.8%; Pred. NO. 2.4e+02; Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 RIIRKIHIIRKY 16
| | : | | : | | |
DB 200 RDINRIIEILKKY 212

RESULT 13

US-08-505-486-65

; Sequence 65, Application US/08505486
; Patent No. 5955573

; GENERAL INFORMATION:

; APPLICANT: Jesse M. Jaynes

; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE

; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND

; NUMBER OF SEQUENCES: 98

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ

; STREET: 555 Thirteenth Street N.W.

; CITY: Washington

; STATE: D. C.

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE

; OPERATING SYSTEM: DOS

; SOFTWARE: Wordperfect 5.1+

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/505,486

;; FILING DATE: 21-JUL-1995
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: U.S. 08/279,472
;; FILING DATE: 22-JUL-1994
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: WALKER, BARBARA W.
;; REGISTRATION NUMBER: 35,400
;; REFERENCE/DOCKET NUMBER: 2093-117A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)783-6040
;; TELEFAX: (202)783-6031
;; INFORMATION FOR SEQ ID NO: 65:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27
;; TYPE: AMINO ACID
;; TOPOLOGY: LINEAR
;; MOLECULE TYPE:
;; DESCRIPTION: PEPTIDE
;; HYPOTHETICAL: NO
;; FRAGMENT TYPE: COMPLETE PEPTIDE
;; ORIGINAL SOURCE: SYNTHETIC
;; IMMEDIATE SOURCE: SYNTHETIC
;; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-505-486-65

Query Match 45.6%; Score 36; DB 2; Length 27;
Best Local Similarity 61.5%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 RRIIRKIIHIKK 15
DB 3 KRIARKILKRIKK 15

RESULT 14
US-08-801-028-65
;; Sequence 65, Application US/08801028
;; Patent No. 6018102
;; GENERAL INFORMATION:
;; APPLICANT: JOAN GARBARINO
;; APPLICANT: JESSE M. JAYNES
;; APPLICANT: WILLIAM BELKNAP
;; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRODUCT
;; NUMBER OF SEQUENCES: 98
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: STEVEN J. HULTQUIST
;; ADDRESS: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
;; STREET: 200 PARK DRIVE, SUITE 210
;; STREET: P.O. BOX 14329
;; CITY: RESEARCH TRIANGLE PARK
;; STATE: NORTH CAROLINA
;; COUNTRY: USA
;; ZIP: 27709
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
;; COMPUTER: APPLE MACINTOSH
;; OPERATING SYSTEM: MACINTOSH
;; SOFTWARE: M.S. WORD 5.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/801,028
;; FILING DATE: 19-FEB-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/279,472
;; FILING DATE: JULY 22, 1994
;; APPLICATION NUMBER: 08/225,476
;; FILING DATE: 04-20-94
;; APPLICATION NUMBER: 08/225,476
;; FILING DATE: 04-08-94
;; APPLICATION NUMBER: 08/039,620
;; FILING DATE: 06-04-93

;; APPLICATION NUMBER: 08/148,491
;; FILING DATE: 11-08-93
;; APPLICATION NUMBER: 08/148,889
;; FILING DATE: 11-08-93
;; ATTORNEY/AGENT INFORMATION:
;; NAME: WASSERMAN, FRAN S.
;; REGISTRATION NUMBER: 34,273
;; REFERENCE/DOCKET NUMBER: 4013-104
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (919)990-9531
;; TELEFAX: (919)990-9532
;; INFORMATION FOR SEQ ID NO: 65:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 27
;; TYPE: AMINO ACID
;; TOPOLOGY: LINEAR
;; MOLECULE TYPE:
;; DESCRIPTION: PEPTIDE
;; HYPOTHETICAL: NO
;; FRAGMENT TYPE: COMPLETE PEPTIDE
;; ORIGINAL SOURCE: SYNTHETIC
;; IMMEDIATE SOURCE: SYNTHETIC
;; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-801-028-65

Query Match 45.6%; Score 36; DB 3; Length 27;
Best Local Similarity 61.5%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 RRIIRKIIHIKK 15
DB 3 KRIARKILKRIKK 15

RESULT 15
US-09-340-154-65
;; Sequence 65, Application US/09340154
;; Patent No. 6084156
;; GENERAL INFORMATION:
;; APPLICANT: Jesse M. Jaynes
;; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
;; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
;; TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
;; NUMBER OF SEQUENCES: 98
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
;; STREET: 555 Thirteenth Street N.W.
;; CITY: Washington
;; STATE: D. C.
;; COUNTRY: USA
;; ZIP: 20004
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
;; COMPUTER: IBM COMPATIBLE
;; OPERATING SYSTEM: DOS
;; SOFTWARE: Wordperfect 5.1+
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/340,154
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/505,486
;; FILING DATE: 21-JUL-1995
;; APPLICATION NUMBER: U.S. 08/279,472
;; FILING DATE: 22-JUL-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: WALKER, BARBARA W.
;; REGISTRATION NUMBER: 35,400
;; REFERENCE/DOCKET NUMBER: 2093-117A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)783-6040
;; TELEFAX: (202)783-6031

```

; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 27
;   TYPE: AMINO ACID
;   TOPOLOGY: LINEAR
; MOLECULE TYPE:
;   DESCRIPTION: PEPTIDE
;   HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
; US-09-340-154-65

```

```

Query Match      45.64; Score 36; DB 3; Length 27;
Best Local Similarity 61.58; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      3 RRIIRKIIHIKK 15
       :|| |||: |||
Db      3 KRIARKILKRIRK 15

```

Search completed: July 12, 2002, 07:55:05
 Job time: 112 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:56:49 ; Search time 95.47 Seconds
(Without alignments)
16.104 Million cell updates/sec

Title: US-09-642-744B-9

Sequence: 1 NRRIRKIIHIKKY 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-71: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	59.5	152	2	S68411 cathelin-related p
2	47	59.5	160	2	S68228 myeloid antimicrob
3	47	59.5	160	2	S68412 cathelin-related p
4	45	57.0	234	2	I40822 sigk protein - C10
5	45	57.0	234	2	D97108 DNA-dependent RNA
6	44	54.4	420	2	S45630 DNA primase chain
7	43	54.4	483	2	S75369 hypothetical prote
8	43	54.4	867	2	G69485 DNA-directed RNA p
9	42	53.2	125	2	A99114 hypothetical prote
10	42	53.2	146	2	D90581 conserved hypotet
11	42	53.2	417	2	A33269 DNA primase (EC 2.
12	42	53.2	537	2	C90389 conserved hypotet
13	41	51.9	186	2	B90116 hypothetical prote
14	41	51.9	337	2	S46010 hypothetical prote
15	41	51.9	507	2	T10753 intestinal DNA rep
16	41	51.9	613	2	F64056 probable ATP-dep
17	40	50.6	105	2	F71128 hypothetical prote
18	40	50.6	221	2	S39609 hypothetical prote
19	40	50.6	221	2	S39609 ERD2 protein - mal
20	40	50.6	370	2	C70464 GTP-binding protei
21	40	50.6	791	2	D82901 ATP-dependent prot
22	40	50.6	1634	2	E64410 DNA-directed DNA p
23	40	50.6	1741	2	T15978 probable phosphori
24	39.5	50.0	208	2	E81935 probable phosphori
25	39.5	50.0	208	2	G81169 N-(5'-phosphoribos
26	39.5	50.0	313	2	A89971 cmp-binding factor
27	39	49.4	41	2	S77768 hypothetical prote
28	39	49.4	113	2	G90223 DNA-directed RNA p
29	39	49.4	151	2	H90035 Tcar transcription

30	39	49.4	202	1	H70318 conserved hypotet
31	39	49.4	210	2	AC1577 precorrin isomerase
32	39	49.4	292	2	A97035 cation efflux syst
33	39	49.4	319	1	S19248 RNA-directed DNA p
34	39	49.4	334	2	B72301 endoglucanase - rh
35	39	49.4	407	2	T37888 hypothetical prote
36	39	49.4	414	2	G64091 cell division prot
37	39	49.4	483	2	B82160 hypothetical prote
38	39	49.4	540	2	T33982 hypothetical prote
39	39	49.4	762	2	S56141 HAK1 protein - yea
40	38.5	48.7	381	1	C64416 conserved hypotet
41	38	48.1	87	2	S00180 spermatic protein
42	38	48.1	96	2	C64354 hypothetical prote
43	38	48.1	156	1	D71689 hypothetical prote
44	38	48.1	172	2	T28288 ORF M5V127 hypothe
45	38	48.1	189	2	H96903 transcription regu

ALIGNMENTS

RESULT 1
S68411
cathelin-related protein 2 precursor - sheep (fragment)
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68411
R:Mahoney, M.M., Lee, A.Y., Brezinski-Caligiuri, D.J., Huttnner, K.M.
FEBS Lett. 377, 519-522, 1995
A:Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobia
A:Reference number: S68411; MUID:96140581
A:Accession: S68411
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1152 <MAH>
A:Cross-references: EMBL:X92757
C:Genetics:
A:Gene: SC5-2
C:Superfamily: cathelin; cystatin homology
F:1-21/Domain: signal sequence #status predicted <SIG>
F:14-122/Domain: cystatin homology <CYS>
F:21-123/Domain: propeptide #status predicted <PRO>
F:124-152/Product: cathelin-related protein 2 #status predicted <MAT>

Query Match 59.5% Score 47; DB 2; Length 152;
Best Local Similarity 60.0% Pred. No. 2.5;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 IRRIRKIIHIKKY 16
DB 126 LRRIGRIAHGVKRY 140

RESULT 2
S68228
myeloid antimicrobial peptide 29 precursor - sheep
N:Alternate names: cathelicidin
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
C:Accession: S68228
R:Bagella, L., Scochi, M., Zanetti, M.
FEBS Lett. 376, 225-228, 1995
A:Title: CDNA sequences of three sheep myeloid cathelicidins.
A:Reference number: S68228; MUID:96105386
A:Accession: S68228
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1160 <BAG>
A:Cross-references: EMBL:L46854; NID:g1161248; PIDN:AAA85470.1; PID:g1161249
C:Superfamily: cathelin; cystatin homology
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-130/Domain: cystatin homology <CYS>

F:29-131/Domain: propeptide #status predicted <PRO>
F:132-160/Product: myeloid antimicrobial peptide 29 #status predicted <MAT>

Query Match

Best Local Similarity 59.5%; Score 47; DB 2; Length 160;
Pred. No. 2.6;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 IRRIRKIIHIKKY 16

Db 134 LRRIGRKIIAHGVKKY 148

RESULT 3

S68412

Cathelin-related protein 1 precursor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000

C:Accession: S68412

R:Mahoney, M.M.; Lee, A.Y.; Brezinski-Caligiuri, D.J.; Hutner, K.M.

FEBS Lett. 377, 519-522, 1995

A:Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobial p

A:Reference number: S68411; MUID:96140581

A:Accession: S68412

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-160 <MAH>

A:Cross-references: EMBL:X92758

C:Genetics:

A:Gene: SC5-1

C:Superfamily: cathelin; cystatin homology

F:1-29/Domain: signal sequence #status predicted <SIG>

F:22-130/Domain: cystatin homology <CTS>

F:30-131/Domain: propeptide #status predicted <PRO>

F:132-160/Product: cathelin-related protein 1 #status predicted <MAT>

Query Match

Best Local Similarity 59.5%; Score 47; DB 2; Length 160;
Pred. No. 2.6;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 IRRIRKIIHIKKY 16

Db 134 LRRIGRKIIAHGVKKY 148

RESULT 4

I40822

sigk protein - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 21-Jul-2000

C:Accession: I40822

R:Sauer, U.; Treuner, A.; Buchholz, M.; Santangelo, J.D.; Durter, P.

J. Bacteriol. 176, 6572-6582, 1994

A:Title: Sporulation and primary sigma factor homologous genes in Clostridium acetobutyli

A:Reference number: I40609; MUID:95050216

A:Accession: I40822

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-234 <RES>

A:Cross-references: GB:L23317; NID:g2745925; PIDN:AAB94775.1; PID:g528974

C:Genetics:

A:Gene: sigk

A:Start codon: GTG

C:Superfamily: transcription initiation factor sigma K; transcription initiation factor

F:56-230/Domain: transcription initiation factor sigma katf homology <KTF>

Query Match

Best Local Similarity 57.0%; Score 45; DB 2; Length 234;
Pred. No. 7.9;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IRRIRKIIHIKKY 16

Db 55 IERNRLVAHIKKY 69

RESULT 5

D97108

DNA-dependent RNA polymerase sigma chain [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: D97108

R:Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L

. J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: D97108

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-234 <KUP>

A:Cross-references: GB:AE001437; PIDN:AAK79655.1; PID:915024653; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1689

C:Superfamily: transcription initiation factor sigma K; transcription initiation fact

Query Match

Best Local Similarity 57.0%; Score 45; DB 2; Length 234;
Pred. No. 7.9;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 IRRIRKIIHIKKY 16

Db 55 IERNRLVAHIKKY 69

RESULT 6

S45630

DNA primase chain p48 - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999

C:Accession: S45630

R:Stadlauer, F.; Brueckner, A.; Rehfuess, C.; Eckerstorn, C.; Lotzpeich, F.; Foerstl

Eur. J. Biochem. 222, 781-793, 1994

A:Title: DNA replication in vitro by recombinant DNA-polymerase-alpha-primase.

A:Reference number: S45628; MUID:94298818

A:Accession: S45630

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-420 <STA>

A:Cross-references: EMBL:X74330; NID:9510405; PIDN:CAAS2377.1; PID:9510406

C:Superfamily: DNA primase 50K chain

Query Match

Best Local Similarity 55.7%; Score 44; DB 2; Length 420;
Pred. No. 20;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 IRRIRKIIHIKKY 16

Db 208 IHPRIKSIHIKKY 222

RESULT 7

S75369

hypochemical protein c04002 - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000

C:Accession: S75369

R:Senen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny,

Mol. Microbiol. 22, 175-191, 1996

A:Title: Organizational characteristics and information content of an archaeal genome

A:Reference number: S73076; MUID:97055432

A:Accession: S75369

A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-483 <SEN>
A:Cross-references: EMBL:Y08257; NID:g1707772; PID:g1707774
A:Experimental source: strain p2
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C:Genetics:
A:Start codon: GTG
C:Superfamily: conserved hypothetical protein M10900

Query Match 54.4%; Score 43; DB 2; Length 483;
Best Local Similarity 69.2%; Pred. No. 33;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 NRRRIIRKIIHII 13
|||||:|:|
Db 378 NRRRIIRSIYEIV 390

RESULT 8
G69485
DNA-directed RNA polymerase, subunit A' (rpoA1) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Jun-1999
C:Accession: G69485
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glocke, A.; Zhou, L.; Oylerbeck, R.; Cocayne, J.D.; Weisman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Ueberback, T.; Cotton, M.D.; Spriggs, T.; Artlisch, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Wiese, C.R.; Ventler, J.C.
A:Title: The complete genome sequence of the hyperthermophilic sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343
A:Accession: G69485
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-867 <KLE>
A:Cross-references: GB:AE000972; GB:AE000782; NID:g2689295; PIDN:AAB89365.1; PID:g264865
C:Superfamily: Halobacterium DNA-directed RNA polymerase chain A

Query Match 54.4%; Score 43; DB 2; Length 867;
Best Local Similarity 53.3%; Pred. No. 57;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NRRRIIRKIIHIKK 15
:|||||:|:|
Db 853 DVRRIRREVGEVK 867

RESULT 9
A99114
hypothetical protein orf125 [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: A99114
R:Douglas, S.; Zanner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671
A:Accession: A99114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <DOU>
A:Cross-references: GB:AJ010592; NID:g12580762; PIDN:CAC27080.1; GSPDB:GN00151
C:Genetics:
A:Gene: orf125
A:Map position: 2
A:Genome: nucleomorph
C:Keywords: nucleomorph

Query Match 53.2%; Score 42; DB 2; Length 125;
Best Local Similarity 70.0%; Pred. No. 13;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 6 IRKIIHIKK 15
|||||:|:|
Db 9 IRKIIHLVRK 18

RESULT 10
D90581
conserved hypothetical protein MYPV_5560 [imported] - Mycoplasma pulmonis (strain UAB
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: D90581
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barde, V.; Sanson, D.; Gallsion, F.; Moser,
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: D90581
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-146 <KUR>
A:Cross-references: GB:AL445566; PID:g14089971; PIDN:CAC13729.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV_5560
A:Genetic code: SGC3
C:Superfamily: Haemophilus influenzae conserved hypothetical protein HI0305

Query Match 53.2%; Score 42; DB 2; Length 146;
Best Local Similarity 43.8%; Pred. No. 15;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 NRRRIIRKIIHIKKY 16
|:|:|:|:|:|
Db 42 NFKKVIINKLKHYEKKY 57

RESULT 11
A33269
DNA primase (EC 2.7.7.-) 46k chain - mouse
N:Alternate names: DNA polymerase alpha/DNA primase complex 46k chain
C:Species: Mus musculus (house mouse)
C:Date: 22-Nov-1989 #sequence_revision 04-Sep-1992 #text_change 17-Nov-2000
C:Accession: A33269; D46642
R:Prussak, C.E.; Almazan, M.T.; Tseng, B.Y.
J. Biol. Chem. 264, 4957-4963, 1989
A:Title: Mouse primase p49 subunit molecular cloning indicates conserved and divergen
A:Reference number: A33269; MUID:89174774
A:Accession: A33269
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-417 <PRU>
A:Cross-references: GB:J04620; NID:g200196; PIDN:AAA39880.1; PID:g200197
A:Note: the authors translated the codon ATC for residue 57 as Tyr
R:Miyaizawa, H.; Izumi, M.; Tada, S.; Takada, R.; Masutani, M.; Ui, M.; Hanaoka, F.
J. Biol. Chem. 268, 8111-8122, 1993
A:Title: Molecular cloning of the CNAs for the four subunits of mouse DNA polymerase
A:Reference number: A46642; MUID:93216788
A:Accession: D46642
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-56, 'Y', 58-417 <MIY>
A:Cross-references: GB:DJ3544; NID:g303660; PIDN:BAA02744.1; PID:g303661
A:Experimental source: FM3A cells
A:Note: sequence extracted from NCBI backbone (NCBIN:129152, NCBIP:129153)
C:Superfamily: DNA primase 50K chain
C:Keywords: nucleotidyltransferase

Fri Jul 12 08:56:29 2002

us-09-642-744b-9.rpr

Page 5

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:04:36 ; Search time 46.15 Seconds

(without alignments)
13.424 Million cell updates/sec

Title: US-09-642-744B-9

Perfect score: 79

Sequence: 1 NRRIRKIHIIKKY 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	59.5	160	SC51_SHEEP	P49928 ovis aries
2	47	59.5	160	SC52_SHEEP	P49929 ovis aries
3	44	55.7	420	PR11_HUMAN	P49642 homo sapien
4	42	53.2	417	PR11_MOUSE	P20664 mus musculu
5	42	53.2	821	MCM6_HUMAN	P14566 homo sapien
6	41	51.9	337	YB21_YEAST	P38276 saccharomyc
7	41	51.9	507	MCM6_RAT	O62724 ratulus norv
8	41	51.9	613	DEAD_HA1IN	P44586 haemophilus
9	40	50.6	105	Y795_PYROH	O58525 pyrococcus
10	40	50.6	221	ERD2_PLAFA	P33948 plasmodium
11	40	50.6	1634	DPOL_METJA	O58295 methanococc
12	39	49.4	414	FTSY_HA1IN	P44870 haemophilus
13	39	49.4	656	V091_F0MPV	O72896 towpox vir
14	39	49.4	821	HAK1_DEBOC	P50505 debaryomyce
15	38.5	48.7	381	TH11_MERJA	O58341 methanococc
16	38	48.1	87	SSSI_SCYCA	P13275 scylliorhinu
17	38	48.1	93	Y435_METJA	O57877 methanococc
18	38	48.1	134	L1PB_HUMAN	O75627 homo sapien
19	38	48.1	138	RRI1_ARATH	P36802 arabidopsi
20	38	48.1	156	Y330_XENTIA	O52438 rickettsia
21	38	48.1	351	CX41_XENTIA	P51914 xenopus lae
22	38	48.1	457	BAG4_HUMAN	O95429 homo sapien
23	38	48.1	464	GATH_YEAST	O03557 saccharomyc
24	38	48.1	552	YB1D_ECOLI	P75828 escherichia
25	38	48.1	623	Y014_MYCPN	P75095 mycoplasma
26	38	48.1	633	YR45_GAEEL	O09552 caenorhabdi
27	38	48.1	670	REP_HA1IN	P44804 haemophilus
28	37	46.8	62	RL30_BACST	P02431 bacillus st
29	37	46.8	100	YAS2_METJA	O58452 methanococc
30	37	46.8	179	CAS2_RAT	P02667 ratulus norv
31	37	46.8	212	ALKH_HA1IN	P44460 h putative
32	37	46.8	333	TAL2_YEAST	P53228 saccharomyc
33	37	46.8	364	LEU3_AOUAE	O66607 aquifex aeo

34	37	46.8	368	1	YSPB_BACSU	P94536 bacillus su
35	37	46.8	399	1	LICH_HUMAN	P38571 homo sapien
36	37	46.8	753	1	YAT9_SCHPO	O09763 schizosacch
37	37	46.8	847	1	ENV_HV1W2	P05880 human immun
38	37	46.8	856	1	ENV_HV1W1	P31872 human immun
39	37	46.8	1229	1	CV43_TRYBB	O99280 trypanosoma
40	36	45.6	67	1	RL30_THEMA	O9x1j1 thermococ
41	36	45.6	108	1	YF24_METJA	O58919 methanococc
42	36	45.6	154	1	YQAR_BACSU	P45914 bacillus su
43	36	45.6	159	1	RS11_SOYBN	P17093 glycine max
44	36	45.6	162	1	IL15_CERAE	P40221 ceropithec
45	36	45.6	162	1	IL15_FELCA	O97687 felis silve

ALIGNMENTS

RESULT	ID	SC51_SHEEP	STANDARD:	PRT:	160 AA.
1	AC	P49928:			
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	01-NOV-1997	(Rel. 35, Last annotation update)			
DE	Cathelin-related peptide SC5 precursor 1 (Antibacterial peptide SMAP-29)				
DE	29) (Myeloid antibacterial peptide SMAP-29).				
OS	Ovis aries (Sheep).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OX	Bovidae; Caprinae; Ovis.				
OX	NCBI_TaxID=9940;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Bone marrow;				
RC	MEDLINE=96140581; PubMed=8549789;				
RA	Maehoney M.M., Lee A.Y., Brezinski-Caliguri D.J., Huttner K.M.;				
RA	"Molecular analysis of the sheep cathelin family reveals a novel				
RT	antimicrobial peptide.";				
RL	FEBS Lett. 377:519-522(1995).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RA	Huttner K.M., Maehoney M.M.;				
RA	Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.				
CC	- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.				
CC	- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: X92757; CAA63412.1; -				
DR	EMBL: U06000; AAB49715.1; -				
DR	InterPro: IPR001894; Cathelicidin.				
DR	Pfam: PF00666; Cathelicidins; 1.				
DR	ProDom: PD001838; Cathelicidins; 1.				
DR	PROSITE: PS00946; Cathelicidins; 1; 1.				
DR	PROSITE: PS00947; CATHELICIDINS_2; 1.				
KW	Antibiotic; Signal.				
FT	SIGNAL	1	29	POTENTIAL.	
FT	PROPEP	30	131	BY SIMILARITY.	
FT	CHAIN	132	160	CATHELIN-RELATED PEPTIDE SC5.	
FT	MOD_RSS	30	30	PYROLIDONE CARBOXYLIC ACID	
FT				(BY SIMILARITY).	
FT	DISULFID	86	97	BY SIMILARITY.	
FT	DISULFID	108	125	BY SIMILARITY.	
FT	SEQUENCE	160 AA;	17786 MW;	BD9B8596432C249 CRC64;	
SQ					

Query Match 59.5%; Score 47; DB 1; Length 160;
 Best Local Similarity 60.0%; Pred. No. 0.63;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 IRRIRKIIHIKKY 16
 DB 134 LRLGRKIIAHGVKKY 148

```

RESULT 2
SCS2_SHEEP STANDARD; PRT; 160 AA.
AC P49929;
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cathelin-related peptide SCS precursor 2 (Antibacterial peptide SMAP-29) (Myeloid antibacterial peptide SMAP-29).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96140581; PubMed=8549789;
RA Mahoney M.M., Lee A.Y., Brezinski-Caliguri D.J., Huttner K.M.;
RT "Molecular analysis of the sheep cathelin family reveals a novel
  antimicrobial peptide."
RL FEBS Lett. 377:519-522(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96105386; PubMed=7498547;
RA Bagella L., Scoochi M., Zanetti M.;
RT "cDNA sequences of three sheep myeloid cathelicidins."
RL FEBS Lett. 376:225-228(1995).
CC -1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
DR EMBL: X92758; CAA63413.1; -
DR EMBL: I46854; AAA85470.1; -
DR InterPro: IPR001894; Cathelicidin.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00666; Cathelicidins; 1.
DR ProDom: PD001838; Cathelicidin; 1.
DR SMART: SM00043; CY; 1.
DR PROSITE: PS00946; CATHELICIDINS_1; 1.
DR PROSITE: PS00947; CATHELICIDINS_2; 1.
KW Antidiotic; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 131 BY SIMILARITY.
FT CHAIN 132 160 CATHELIN-RELATED PEPTIDE SCS.
FT MOD_RES 30 30 PYROLIDONE CARBOXYLIC ACID
  (BY SIMILARITY).
FT DISULFID 86 97 BY SIMILARITY.
FT DISULFID 108 125 BY SIMILARITY.
FT CONFLICT 28 28 S -> R (IN REF. 2).
FT SEQUENCE 160 AA; 17742 MW; 4FB98A09355B51F CRC64;

```

Query Match 59.5%; Score 47; DB 1; Length 160;
 Best Local Similarity 60.0%; Pred. No. 0.63;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 IRRIRKIIHIKKY 16
 DB 134 LRLGRKIIAHGVKKY 148

```

RESULT 3
PRIL_HUMAN STANDARD; PRT; 420 AA.
ID PRIL_HUMAN
AC P49642;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA primase small subunit (EC 2.7.7.-) (DNA primase 49 kDa subunit)
DE (p49).
GN PRIM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94298818; PubMed=8026492;
RA Stadbauer F., Brueckner A., Reifuss C., Eckerskorn C.,
RA Lottspeich F., Foerster V., Tseng B.Y., Nashner H.P.;
RT "DNA replication in vitro by recombinant DNA-polymerase-alpha-
  primase".
RL Eur. J. Biochem. 222:781-793(1994).
RN [2]
RP SEQUENCE OF 97-146 FROM N.A.
RX MEDLINE=97422622; PubMed=9266648;
RA Cloutier S., Hamel H., Champagne M., Yotov W.V.;
RT "Mapping of the human DNA primase 1 (PRIM1) to chromosome 12q13."
RL Genomics 43:398-401(1997).
CC -1- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
  RNA PRIMERS FOR THE OKAZAKI FRAGMENTS MADE DURING DISCONTINUOUS
  DNA REPLICATION.
CC -1- SUBUNIT: HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC PRIMASE SMALL SUBUNIT
  FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
DR EMBL: X74330; CAA52377.1; -
DR EMBL: U89689; AAC51726.1; -
DR MIM: 176635; -
DR InterPro: IPR002755; DNA_Primase-S.
DR Pfam: PF01896; DNA_Primase-S; 1.
DR Transferrase; DNA replication; DNA-directed RNA polymerase; Primosome.
KW ACT_SITE 44 44 POTENTIAL.
FT ACT_SITE 109 109 POTENTIAL.
FT ACT_SITE 111 111 POTENTIAL.
FT METAL 121 121 POTENTIAL.
FT METAL 122 122 POTENTIAL.
FT METAL 128 128 POTENTIAL.
FT METAL 131 131 POTENTIAL.
FT SEQUENCE 420 AA; 49902 MW; 9B5AC900E0C3CCE8 CRC64;

```

Query Match 55.7%; Score 44; DB 1; Length 420;
 Best Local Similarity 66.7%; Pred. No. 5.1;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 IRRIRKIIHIKKY 16
 DB 208 IHPFRKSINIKKY 222

```

RESULT 4
PR1_MOUSE
ID PR1_MOUSE STANDARD: PRT: 417 AA.
AC P20664:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE DNA primase small subunit (EC 2.7.7.-) (DNA primase 49 kDa subunit)
DE (P49).
GN PRIM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Lymphoid;
RX MEDLINE=89174774; PubMed=2925677;
RA Prussak C.E., Almazan M.T., Tseng B.Y.,
RT "Mouse primase p49 subunit molecular cloning indicates conserved and
RT divergent regions.";
RL J. Biol. Chem. 264:4957-4963(1989).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 51-66 AND 222-237.
RX MEDLINE=93216788; PubMed=8463324;
RA Miyazawa H., Izumi M., Tada S., Takada R., Masutani M., Ui M.,
RA Hanoka F.;
RT "Molecular cloning of the cDNAs for the four subunits of mouse DNA
RT polymerase alpha-primase complex and their gene expression during
RT cell proliferation and the cell cycle.";
RL J. Biol. Chem. 268:8111-8122(1993).
CC -1- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
CC RNA PRIMERS FOR THE OKAZAKI FRAGMENTS MADE DURING DISCONTINUOUS
CC DNA REPLICATION
CC -1- SUBUNIT: HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC PRIMASE SMALL SUBUNIT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J04620; AAA39880.1; -
DR EMBL: D13544; BAA02744.1; -
DR PIR: A33269; A33269.
DR MGD: MGI:97757; Prim1.
DR InterPro: IPR002755; DNA_primase_S.
KW Pfam: PF01896; DNA_primase_S.
DR Transferrase; DNA replication; DNA-directed RNA polymerase; Primosome.
FT ACT_SITE 44 44
FT ACT_SITE 109 109
FT ACT_SITE 111 111
FT METAL 121 121
FT METAL 122 122
FT METAL 128 128
FT METAL 131 131
FT METAL 131 131
FT CONFLICT 57 57
FT CONFLICT 417 AA; 49295 MW; 54BFB4DADE47D8A CRC64;
SQ SEQUENCE 417 AA; 49295 MW; 54BFB4DADE47D8A CRC64;

Query Match 53.2%; Score 42; DB 1; Length 417;
Best Local Similarity 53.3%; Pred. No. 11;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 2 IRRIRKIHIIKKY 16
DB 208 VHPFVKRSINIIRKY 222

```

```

RESULT 5
MCM6_HUMAN
ID MCM6_HUMAN STANDARD: PRT: 821 AA.
AC Q14566; Q13504; Q99859;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA replication licensing factor MCM6 (P105MCM).
DE MCM6.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=97431351; PubMed=9286856;
RA Tsunaga H., Yabuta N., Hosoya S., Tamura K., Endo Y., Nojima H.;
RT "hSMCM6: a new member of the human MCM/PI family encodes a protein
RT homologous to fission yeast Mts5.";
RL Genes Cells 2:381-399(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98184833; PubMed=9516426;
RA Holthoff H.P., Baack M., Richter A., Ritzel M., Kuipers R.;
RT "Human protein MCM6 on HeLa cell chromatin.";
RL J. Biol. Chem. 273:7320-7325(1998).
RN [3]
RP SEQUENCE OF 640-821 FROM N.A.
RX MEDLINE=97131582; PubMed=8977093;
RA Harvey C.B., Wang Y., Darmoul D., Phillips A., Mantel N.,
RA Swallow D.M.;
RT "Characterisation of a human homologue of a yeast cell division cycle
RT gene, MCM6, located adjacent to the 5' end of the lactase gene on
RT chromosome 2q21.";
RL FEBS Lett. 398:135-135(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF A SINGLE ROUND OF DNA
CC REPLICATION DURING S PHASE. BINDS TO CHROMATIN DURING G1 AND
CC DETACH FROM IT DURING S PHASE AS IF IT LICENSES THE CHROMATIN TO
CC REPLICATE.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE MCM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D94557; BAA12699.1; -
DR EMBL: U46836; AAC50766.1; -
DR EMBL: U67284; AAB48165.1; -
DR EMBL: U67281; AAB48165.1; JOINED.
DR EMBL: U67282; AAB48165.1; JOINED.
DR EMBL: U67283; AAB48165.1; JOINED.
DR MIM: 601806; -
DR InterPro: IPR001208; MCM.
DR Pfam: PF00493; MCM; 1.
DR SMART: SM00350; MCM; 1.
DR PROSITE: PS00847; MCM_1; 1.
DR PROSITE: PS0051; MCM_2; 1.
KW DNA replication regulation; DNA-binding; Nuclear protein;
KW DNA replication; ATP-binding; Cell cycle.
FT NP_BIND 346 553
FT DOMAIN 366 403
FT CONFLICT 377 387
FT CONFLICT 495 495
FT CONFLICT 738 738
FT CONFLICT 790 790
FT CONFLICT 821 AA; 92869 MW; F94496BEB25A3E501 CRC64;
SQ SEQUENCE 821 AA; 92869 MW; F94496BEB25A3E501 CRC64;

```

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirness E.F.,
RA Kurland A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

```
RA Scott J.D., Shirley R., Liu L.-I., Glodex A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sadek D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Georgiagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: HAS A HELIX-DESTABILIZING ACTIVITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32709; AAC21900.1; -.
DR HSSP: Q58083; 1HV8.
DR TIGR: H10231; -.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000629; DEAD_ATP_helicase.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF00270; DEAD. 1.
DR Pfam: PF00271; Helicase_C. 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELICG; 1.
DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.
KM Hydrolyase: Helicase; ATP-binding; RNA-binding;
KM Transcription regulation: Complete proteome.
FT NP_BIND 49 56 ATP (POTENTIAL).
FT SITE 155 158 DEAD BOX.
SQ SEQUENCE 613 AA; 69705 MW; 1B826CBDEB1704DF CRC64;
```

```
CC -1- SIMILARITY: BELONGS TO THE UPF0148 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000003; BAA29888.1; -.
KM Hypothetical protein: Complete proteome.
SQ SEQUENCE 105 AA; 12286 MW; B55443FACC88F376 CRC64;
```

Query Match
Best Local Similarity 50.6%; Score 40; DB 1; Length 105;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
QY 4 RIIRKIHIKKY 16
DB 86 RYMEKIIEVLEKY 98

RESULT 10
ID ERD2_PLAFA STANDARD: PRT; 221 AA.
AC P33948;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE ER lumen protein retaining receptor.
CN ERD2.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OC NCBL_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ITGPF6.
RX MEDLINE-94038960; PubMed-8223485;
RA Elmsdorf H.G., Halder K.;
RT "Identification and localization of ERD2 in the malaria parasite
RT Plasmodium falciparum: separation from sites of sphingomyelin
RT synthesis and implications for organization of the Golgi."
RL EMBL J.124763-4773(1993).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ERD2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74869; CAAS2861.1; -.
DR EMBL: Z26043; CAA81128.1; -.
DR PIR: S42742; S42742.
DR PIR: S39609; S39609.
DR InterPro: IPR000133; ER_lumen_receptor.
DR Pfam: PF00810; ER_lumen_receptor.
DR PRINTS: PR00660; ERLUMENR.
DR ProDom: PD005774; ER_lumen_receptor; 1.
DR PROSITE: PS00951; ER_LUMEN_RECEPTOR_1; 1.
DR PROSITE: PS00952; ER_LUMEN_RECEPTOR_2; 1.
KW Endoplasmic reticulum: Transmembrane; Protein transport: Receptor.
FT DOMAIN 1 2 LUMENAL (POTENTIAL).
FT TRANSMEM 3 21 POTENTIAL.
FT DOMAIN 22 35 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 36 52 POTENTIAL.
FT DOMAIN 53 61 LUMENAL (POTENTIAL).
FT TRANSMEM 62 80 POTENTIAL.
```



```

FT DOMAIN 81 99 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 100 113 POTENTIAL.
FT DOMAIN 114 120 LUMENAL (POTENTIAL).
FT TRANSMEM 121 140 POTENTIAL.
FT TRANSMEM 141 152 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 153 171 POTENTIAL.
FT DOMAIN 172 183 LUMENAL (POTENTIAL).
FT TRANSMEM 184 203 POTENTIAL.
FT DOMAIN 204 221 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 221 AA; 26446 MW; 7F231ED6AAACB368 CRC64;

Query Match 50.6%; Score 40; DB 1; Length 221;
Best Local Similarity 43.8%; Pred. No. 12;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 NRRIRIKIHIIRKY 16
   | | | | | | | | | |
Db 2 NIFRLIGDILHVSMTY 17

RESULT 11
DPOL_METJA STANDARD; PRT; 1634 AA.
AC 058295;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) [Contains: Mja pol-1 intein; Mja pol-2
  intein].
GN POL OR Mj0885.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota; Methanococcales; Methanococcaceae;
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=9633799; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
  Koutson G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
  Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
  Overbeek R., Kirkness E.F., Weissstock K.G., Merrick J.M., Glodek A.,
  Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
  Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
  Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
  Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
  jannaschii."
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
  + (DNA)(N).
CC -1- PFM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
  A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
  FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67532; AAB8889.1; -.
CC HSSP: P56689; ITGO.
CC TIGR: Mj0885; -.
CC InterPro: IPR002064; DNA_pol_B.
CC InterPro: IPR003386; HIntC.
CC InterPro: IPR003587; HIntC.
CC InterPro: IPR002203; InteIn.
CC InterPro: IPR004042; InteIn_endonuc.

```

```

DR Pfam: PF00136; DNA_pol_B: 3.
DR PRINTS: PF00106; DNAPOLB.
DR SMART: SM00305; HIntC: 2.
DR SMART: SM00306; HIntC: 2.
DR SMART: SM00486; POLB: 1.
DR PROSITE: PS00116; DNA_POLYMERASE_B: 1.
DR PROSITE: PS00881; PROTEIN_SPLICING: 2.
DR Transferrase: DNA-directed DNA polymerase; DNA replication;
  KW DNA-binding; Autocatalytic cleavage; Protein splicing;
  KW Complete proteome.
FT CHAIN 1 425 POL, 1ST PART (POTENTIAL).
FT CHAIN 426 794 MJA POL-1 INTEIN (POTENTIAL).
FT CHAIN 795 882 POL, 2ND PART (POTENTIAL).
FT CHAIN 883 1358 MJA POL-2 INTEIN (POTENTIAL).
FT CHAIN 1359 1634 POL, 3RD PART (POTENTIAL).
SQ SEQUENCE 1634 AA; 191708 MW; 8A1FAFAB1F97DD CRC64;

Query Match 50.6%; Score 40; DB 1; Length 1634;
Best Local Similarity 43.8%; Pred. No. 91;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 1 NRRIRIKIHIIRKY 16
   | | | | | | | | | |
Db 208 NEKLRIKIETKEY 223

RESULT 12
FTSY_HAEIN STANDARD; PRT; 414 AA.
AC P4870;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cell division protein ftsy homolog.
GN FTSY OR HI0768.
OS Haemophilus influenzae.
OC Bacteria: Proteobacteria; gamma subdivision; Pasteurellaceae;
  Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
  Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
  Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
  Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
  Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
  Uterback T.R., Hanna M.C., Nguyen D.T., Sautek D.M., Brandon R.C.,
  Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
  Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
  Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
  influenzae Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: FUNCTIONAL HOMOLOGY OF SRP RECEPTOR. PROBABLY INVOLVED IN
  THE RECEPTION AND INSERTION OF A SUBSET OF PROTEINS AT THE
  CYTOPLASMIC MEMBRANE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32760; AAC22426.1; -.
CC HSSP: P10121; LTS.

```

DR TIGR: H10768; -
DR InterPro: IPR000897; SRP54.
DR Pfam: PF00448; SRP54; 1.
DR Pfam: PF02881; SRP54.N; 1.
DR PRODOM: PD000819; SRP54; 1.
DR PROSITE: PS00300; SRP54; 1.
KW Signal recognition particle; GTP-binding; RNA-binding; Membrane;
KW Inner membrane; Cell division; Complete proteome.
FT NP_BIND 216 223 GTP (BY SIMILARITY).
FT NP_BIND 298 302 GTP (BY SIMILARITY).
FT NP_BIND 362 365 GTP (BY SIMILARITY).
SQ SEQUENCE 414 AA; 46146 MW; DB2C457C651C52C0 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 414;
Best Local Similarity 37.5%; Pred. No. 33;
Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 NRRIRKIIHIKKY 16
DB 308 NLMDLKKIIVMKKY 323

RESULT 13
V091_FOWPV STANDARD; PRT; 656 AA.
ID V091_FOWPV
AC 072896;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein FPOV91.
GN FPOV91 OR FPOVL.
OS Fowlpox virus (FPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
NC NCB1_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PP-9 / Isolate HP-440;
RX MEDLINE-98451804; PubMed-9778782;
RA Pollitt E., Skinner M.A., Heaphy S.;
RT "Nucleotide sequence of the 4.3 kbp BamHI-N fragment of fowlpox virus
FPV.";
RL Virus Genes 17:5-9(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-20193820; PubMed-10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -1- SIMILARITY: BELONGS TO THE POXYVIRUSES OI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ223385; CA11289.1; -
DR EMBL: AF198100; AAF4435.1; -
SQ SEQUENCE 656 AA; 77177 MW; 274E70D4E26A6456 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 656;
Best Local Similarity 63.6%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 6 IRKIIHIKKY 16
DB 199 VDKIIVIRKY 209

RESULT 14
HAK1_DEBOC STANDARD; PRT; 821 AA.
ID HAK1_DEBOC
AC P50505;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE High affinity potassium transporter.
GN HAK1.
OS Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCB1_TaxID=27300;
RN [1]
RP SEQUENCE FROM N.A.
RA Rodriguez-Navarro A.;
RL submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 60-821 FROM N.A.
RX MEDLINE-95347328; PubMed=7621817;
RA Bannuelos M.A., Klein R.D., Alexander-Bowman S.J.,
RA Rodriguez-Navarro A.;
RT "A potassium transporter of the yeast Schwanniomyces occidentalis
RT homologous to the Kup system of Escherichia coli has a high
RT concentrative capacity.";
RL EMBL J. 14:3021-3027(1995).

CC -1- FUNCTION: MAJOR HIGH-AFFINITY POTASSIUM UPTAKE PROTEIN.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: TO E.COLI POTASSIUM TRANSPORT KUP.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: U22945; AB017122.2; -
DR InterPro: IPR003855; K+-transport.
DR Pfam: PF02705; K_trans; 1.
KW Transmembrane; Transport; Potassium transport.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 191 211 POTENTIAL.
FT TRANSMEM 261 281 POTENTIAL.
FT TRANSMEM 341 361 POTENTIAL.
FT TRANSMEM 387 407 POTENTIAL.
FT TRANSMEM 435 455 POTENTIAL.
FT TRANSMEM 464 484 POTENTIAL.
FT TRANSMEM 492 512 POTENTIAL.
FT TRANSMEM 517 537 POTENTIAL.
SQ SEQUENCE 821 AA; 91935 MW; ECAEA760603F63C2 CRC64;

Query Match 49.4%; Score 39; DB 1; Length 821;
Best Local Similarity 77.8%; Pred. No. 66;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 KIIRIIRKY 16
DB 423 KIIRVSKKY 431

RESULT 15
THII_METUA STANDARD; PRT; 381 AA.
ID THII_METUA
AC 058341;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable thiamine biosynthesis protein thII.
THII OR MJ0931.

```

OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_Taxid=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Kleink H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.
RL Science 273:1058-1073(1996).
CC -!- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE THIAZOLE MOIETY (BY
CC SIMILARITY).
CC -!- PATHWAY: THIAMINE BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: BELONGS TO THE THIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U67536; AAB98933.1; -.
DR TIGR: MJ0931; -.
DR InterPro: IPR004114; THUMP.
DR InterPro: IPR003720; Thit.
DR Pfam: PF02368; Thit; 1.
DR Pfam: PF02926; THUMP; 1.
KW Thiamine biosynthesis; Complete proteome.
SQ SEQUENCE 381 AA; 43436 MW; 0A31F1069DA3357B CRC64;

```

```

Query Match          48.7%; Score 38.5; DB 1; Length 381;
Best Local Similarity 47.1%; Pred. No. 37;
Matches 8; Conservative 5; Mismatches 3; Indels 1; Gaps 1;
QY 1 NIKRIIRK-IHIIKKY 16
   | : | | | : | |
DB 28 NLEELIKKNIKKLRKY 44

```

Search completed: July 12, 2002, 08:04:37
 Job time: 649 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:03:44 ; Search time 174.7 Seconds
(Without alignments)
15.844 Million cell updates/sec

Title: US-09-642-744B-9
Perfect score: 79
Sequence: 1 NRRIRKRIHIKKY 16

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17294929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP-REMBL_19:*
2: sp-archaea:*
3: sp-bacteria:*
4: sp-fungi:*
5: sp-human:*
6: sp-invertebrate:*
7: sp-mammal:*
8: sp-mhc:*
9: sp-organelle:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	60.8	500	3 Q9P6R0	Q9P6R0 schizosacch
2	45	57.0	234	16 Q59264	Q59264 clostridium
3	44	55.7	233	3 Q9L7X2	Q9L7X2 clostridium
4	44	55.7	460	5 Q9VLO3	Q9VLO3 drosophila
5	43	54.4	483	17 P95941	P95941 sulfolobus
6	43	54.4	867	17 Q28391	Q28391 archaeoglob
7	42	53.2	125	10 Q9AVZ1	Q9AVZ1 guillardi
8	42	53.2	146	16 Q9BQ13	Q9BQ13 mycoplasma
9	42	53.2	345	11 Q9D0G7	Q9D0G7 mus musculu
10	42	53.2	414	10 Q947U2	Q947U2 oryza sativ
11	42	53.2	415	11 Q89045	Q89045 ratius notv
12	42	53.2	530	3 Q9C2F6	Q9C2F6 neurospora
13	42	53.2	537	17 Q97W16	Q97W16 sulfolobus
14	42	53.2	954	5 Q9G015	Q9G015 giardia lam
15	42	53.2	969	5 Q9U022	Q9U022 giardia lam
16	41	51.9	146	2 Q9A1H3	Q9A1H3 candidatus

17	41	51.9	152	8 Q02220	Q02220 phytophthor
18	41	51.9	152	8 Q35523	Q35523 phytophthor
19	41	51.9	186	10 Q9AVW6	Q9AVW6 guillardi
20	41	51.9	840	12 Q9DHN4	Q9DHN4 yaba-like d
21	40	50.6	104	2 Q45020	Q45020 borrelia bu
22	40	50.6	106	2 Q44868	Q44868 borrelia bu
23	40	50.6	177	12 Q9YVP1	Q9YVP1 melanoplus
24	40	50.6	234	5 Q952W1	Q952W1 caenorhabdi
25	40	50.6	290	2 Q9S0H6	Q9S0H6 borrelia bu
26	40	50.6	309	2 Q44823	Q44823 borrelia bu
27	40	50.6	337	2 Q9S0U7	Q9S0U7 borrelia bu
28	40	50.6	352	2 Q9S082	Q9S082 borrelia bu
29	40	50.6	357	2 Q9S068	Q9S068 borrelia bu
30	40	50.6	370	16 Q67743	Q67743 aquifex aeo
31	40	50.6	386	12 Q9DHR5	Q9DHR5 yaba-like d
32	40	50.6	438	12 Q91SH5	Q91SH5 guinea pig
33	40	50.6	608	10 Q941Y3	Q941Y3 oryza sativ
34	40	50.6	636	12 Q90142	Q90142 spodoptera
35	40	50.6	791	16 Q9P0E3	Q9P0E3 ureaplasma
36	40	50.6	1063	12 Q9J844	Q9J844 spodoptera
37	39.5	50.0	208	16 Q9K0C6	Q9K0C6 neisseria m
38	39.5	50.0	208	16 Q9UYD1	Q9UYD1 neisseria m
39	39.5	50.0	313	16 Q53606	Q53606 staphylococ
40	39	49.4	41	2 Q48971	Q48971 mycoplasma
41	39	49.4	113	17 Q9UXD9	Q9UXD9 sulfolobus
42	39	49.4	151	16 Q9F4G3	Q9F4G3 staphylococ
43	39	49.4	202	16 Q66580	Q66580 aquifex aeo
44	39	49.4	210	16 Q92CL5	Q92CL5 listeria in
45	39	49.4	216	11 Q9CUL8	Q9CUL8 mus musculu

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	500 AA.
Q9P6R0	AC	Q9P6R0:		
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	TRANSCRIPTION FACTOR IIB 70 KDA SUBUNIT.			
GN	SPB13E7.10C.			
OS	Schizosaccharomyces pombe (fission yeast).			
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;			
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;			
OC	Schizosaccharomyces.			
OX	NCBI_TaxID=4896;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=972H-;			
RA	Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL354632; CAB89885.1; -			
DR	InterPro: IPR004366; Cytlin.			
DR	InterPro: IPR000812; TFIIB.			
DR	Pfam: PF00382; transcript1fac2; 2.			
DR	PRINTS; PR00685; TIFACTOR1B.			
DR	SMART; SM00385; CYCLIN; 2.			
SQ	SEQUENCE 500 AA; 56749 MW; 6CFLFD9ADA2B1B2 CRC64;			

Query Match 60.8%; Score 48; DB 3; Length 500;
Best Local Similarity 42.9%; Pred. No. 11;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 NRRIRKRIHIK 14
DB 245 NRRSVREVVHYK 258

RESULT 2 Q59264

```

ID 059264      PRELIMINARY;      PRT;      234 AA.
AC 059264;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE RNA POLYMERASE SIGMA FACTOR (DNA-DEPENDENT RNA POLYMERASE SIGMA
DE SUBUNIT).
GN SIGK OR CAC1689.
OS Clostridium acetobutylicum.
OC Bacteria: Firmicutes; Bacillus/Clostridium group: Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95050216; PubMed=7961408;
RA Sauer U., Treuner A., Buchholz M., Santangelo J.D., Durre P.;
RT "Sporeulation and primary sigma factor homologous genes in Clostridium
RT acetobutylicum.";
RL J. Bacteriol. 176:6572-6582(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RC MEDLINE=21359325; PubMed=1146286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED.
CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
CC
CC EMBL: I23317; AAB94775.1; -.
DR EMBL; AE007679; AAK79655.1; -.
DR HSSP; P00579; 1SIG.
DR InterPro: IPR000943; Sigma_70.
DR Pfam: PR00140; sigma70_1; 1.
DR PROSITE; PS00715; SIGMA70_1; 1.
DR PROSITE; PS00716; SIGMA70_2; UNKNOWN_1.
KW DNA-binding; DNA-directed RNA polymerase; Sigma factor; Sporulation;
KW Transcription regulation; Complete proteome.
SQ SEQUENCE 234 AA; 26586 MW; 091DD27A22BC1E4 CRC64;

```

```

Query Match      57.0%; Score 45; DB 16; Length 234;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 2 IRRIRKIIHIKKY 16
Db 55 IERNLRVAHVKKY 69

RESULT 3
O9L7X2      PRELIMINARY;      PRT;      233 AA.
AC 09L7X2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE RNA POLYMERASE SIGMA FACTOR.
GN SIGK.
OS Clostridium perfringens.
OC Bacteria: Firmicutes; Bacillus/Clostridium group: Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 8798;
RA Stirewalt V.L., Melville S.B.;
RT "Cloning and sequence analysis of the sigK gene of Clostridium

```

```

RT perfringens.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
CC
CC EMBL: AF218835; AAF26431.1; -.
DR EMBL; AF218835; AAF26431.1; -.
DR HSSP; P00579; 1SIG.
DR InterPro: IPR000943; Sigma_70.
DR Pfam; PF00140; sigma70_1.
DR PRINTS; PR00046; SIGMA70PCT.
DR PROSITE; PS00715; SIGMA70_1; 1.
DR PROSITE; PS00716; SIGMA70_2; 1.
KW DNA-binding; DNA-directed RNA polymerase; Sigma factor;
KW Transcription regulation.
SQ SEQUENCE 233 AA; 26573 MW; 934DA5B89490C1B4 CRC64;

```

```

Query Match      55.7%; Score 44; DB 2; Length 233;
Best Local Similarity 46.7%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 2 IRRIRKIIHIKKY 16
Db 55 VERNLRVAHVKKY 69

RESULT 4
O9VL03      PRELIMINARY;      PRT;      460 AA.
AC 09VL03;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE CG3138 PROTEIN.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Adair J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman J.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland J.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003628; AAF52902.1;
 DR FlyBase: FBgn0032211; CG3138.
 SQ SEQUENCE 460 AA; 53135 MW; C038EC3C799814C3 CRC64;

Query Match 55.7%; Score 44; DB 5; Length 460;
 Best Local Similarity 46.7%; Pred. No. 42;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 IRRIRKIIHIIRKY 16
 Db 161 VERIKIKIVHPKPY 175

RESULT 5
 P95941 PRELIMINARY; PRT; 483 AA.
 AC P95941;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE VIRB HOMOLOG (TYPE II SECRETION SYSTEM PROTEIN, VIRB HOMOLOG)
 DE (GSPF-1).
 GN GSPF-1.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aveyez M.V., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Regan M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL: Y08257; CAA69572.1;
 DR EMBL: AE006651; AAK40475.1;
 DR InterPro: IPR001687; ATP_GTP_A.
 DR InterPro: IPR001482; GSP11_E.
 DR Pfam: PF00437; GSP11_E; 1.
 DR Complete proteome.
 SQ SEQUENCE 483 AA; 55370 MW; C7FC726CA28B7ECF CRC64;

Query Match 54.4%; Score 43; DB 17; Length 483;
 Best Local Similarity 69.2%; Pred. No. 64;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 NRRIRKIIHIIT 13
 Db 378 NRRIRSIYEIV 390
 RESULT 6
 ID 028391 PRELIMINARY; PRT; 867 AA.
 AC 028391;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE, SUBUNIT A' (RPOA1).
 GN AFI188.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uutterback T.,
 RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaean *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 DR EMBL: AE000972; AAB89365.1; -.
 DR TIGR: AFI188;
 DR InterPro: IPR00722; RNA-pol_A.
 DR Pfam: PF00623; RNA-pol_A; 1.
 KW Hypothetical protein; DNA-directed RNA polymerase; Complete proteome.
 SQ SEQUENCE 867 AA; 97806 MW; 6D80511133210085 CRC64;

Query Match 54.4%; Score 43; DB 17; Length 867;
 Best Local Similarity 53.3%; Pred. No. 1,1e+02;
 Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NRRIRKIIHIITK 15
 Db 853 DVRRIRREIVGEVKK 867

RESULT 7
 ID 09AVZ1 PRELIMINARY; PRT; 125 AA.
 AC 09AVZ1;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 15.2 KDA PROTEIN.
 OS Guillardia theta (*Cryptomonas phi*).
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 OX NCBI_TaxID=55529;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20087226; PubMed=10618395;
 RA Zauner S., Fraunholz M., Wastl J., Penny S.L., Beaton M.,
 RA Cavalier-Smith T., Maier U., Douglas S.;
 RT "Chloroplast protein and centosomal genes, a tRNA intron, and odd
 RT telomeres in an unusually compact eukaryotic genome, the cryptomonad
 RT nucleomorph.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21223349; PubMed=11323671;
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.,
 RA Wu X., Reich M., Cavalier-Smith T., Maier U.;
 RT "The highly reduced genome of an enslaved algal nucleus.";
 RL Nature 410:1091-1096(2001).
 DR EMBL: AJ010592; CAC27080.1; -.
 DR Hypothetical protein.
 SQ SEQUENCE 125 AA; 15207 MW; AAD538B5593997A1 CRC64;

```

Query Match          53.2%; Score 42; DB 10; Length 125;
Best Local Similarity 70.0%; Pred. No. 27;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 6 IRRIIRKIIHIKK 15
    |||||:::|
DB 9 IRKIIHLVRK 18

RESULT 8
O98013 PRELIMINARY; PRT; 146 AA.
AC O98013;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DE HYPOTHEICAL PROTEIN MYPU_5560.
GN MYPU_5560.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Hellig R., Ferris S., Barbe V., Samson D., Gallison F.,
RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis."
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL: AL445565; CAC13729.1; -.
DR Hypulist: MYPU_5560; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 146 AA; 16929 MW; 3874101E7539CCF6 CRC64;

Query Match          53.2%; Score 42; DB 16; Length 146;
Best Local Similarity 43.8%; Pred. No. 31;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 NRIIRKIIHIKKY 16
    |:::|:::|
DB 42 NFKVINKLKHKKY 57

RESULT 9
O990G7 PRELIMINARY; PRT; 345 AA.
AC O990G7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE 10 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2610019C06, FULL INSERT SEQUENCE.
GN PRIM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gasterland T., Gissi C., King B., Kochiwa H.,

```

```

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK011456; BAB27629.1; -.
DR MGD: MGI:97757; Prim1.
DR InterPro: IPR002755; DNA_primase-S.
DR Pfam: PF01896; DNA_primase-S; 1.
SQ SEQUENCE 345 AA; 40250 MW; B1ABFB7314D6E1 CRC64;

Query Match          53.2%; Score 42; DB 11; Length 345;
Best Local Similarity 53.3%; Pred. No. 67;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 IRIIRKIIHIKKY 16
    :|||:|||||
DB 136 VHPVRKSNIIKKY 150

RESULT 10
O94702 PRELIMINARY; PRT; 414 AA.
AC O94702;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PUTATIVE TRANSLATION INITIATION FACTOR 2B BETA SUBUNIT.
GN OSJNBAD068A07.14.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RA de la Bastide M., Spiegel L., Kirchoff K., Preston R., King L.,
RA Vil M.D., Baker J., Zuberlarn T., Santos L., Miller B., Kilt K.,
RA Connus D.M., Bell M., Balija V., Shah R., Bahret A., Dike S.,
RA Yang C., O'Shaughnessy A., Palmer L., Dedhia N., McCombie W.R.;
RT "Genomic Sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNBAD068A07, from Chromosome 10, complete sequence."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC091734; AAK98757.1; -.
KM Initiation factor.
SQ SEQUENCE 414 AA; 44646 MW; BA54863017C7EFOC CRC64;

Query Match          53.2%; Score 42; DB 10; Length 414;
Best Local Similarity 42.9%; Pred. No. 80;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 2 IRIIRKIIHIKK 15
    :||:::||||:
DB 73 VGNIVRVLHIITKE 86

RESULT 11
O89045 PRELIMINARY; PRT; 415 AA.
AC O89045;
DT 01-NOV-1998 (TReMBLrel. 08, Created)

```

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE DNA POLYMERASE ALPHA SUBUNIT IV (PRIMASE) (EC 2.7.7.7)
 DE (FRAGMENT).
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE DAWLEY; TISSUE=LIVER;
 RA Popanda O., Flohr C., Thielmann H.W.;
 RT "A mutation in the gene of subunit II of DNA polymerase alpha from
 RT Novikoff cells is concomitant with altered physico-chemical properties
 RT of the enzyme."
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ011608; CA09723.1; -;
 DR InterPro: IPR002755; DNA_primase_5.
 DR Pfam: PF01896; DNA_primase_5; 1.
 KW Transferase; Nucleotidyltransferase.
 FT NON_TER 1
 FT 415 415
 SO SEQUENCE 415 AA; 49140 MW; A3342B9B70FDD2A CRC64;

Query Match 53.2%; Score 42; DB 11; Length 415;
 Best Local Similarity 53.3%; Pred. No. 80;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 NRRIRKIIHIKKY 16
 : :|||:|||||
 Db 210 VHPFVKRSIIKKY 224

RESULT 12
 ID Q9C2F6 PRELIMINARY; PRT; 530 AA.
 AC Q9C2F6;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE RELATED TO GTPASE MSS1, MITOCHONDRIAL.
 GN 9G6.080.
 GN Neurospora crassa.
 OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Sordariales; Sordariaceae; Neurospora.
 CC NCBI_TaxID=5141;
 RN
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
 RA Nykatura G., Meves H.W., Mannhaupt G.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL513463; CAC28756.2; -;
 SO SEQUENCE 530 AA; 56936 MW; 269830E07A4C055 CRC64;

Query Match 53.2%; Score 42; DB 3; Length 530;
 Best Local Similarity 42.9%; Pred. No. 1e+02;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 NRRIRKIIHIKK 14
 : :|:|:|:|:|
 Db 196 NVTRLVKELHSIK 209

RESULT 13
 ID Q97WL6 PRELIMINARY; PRT; 537 AA.
 AC Q97WL6;
 OY 1 NRRIRKIIHIKK 13
 : :|:|:|:|:|
 Db 368 NLRNVLRFVPHIL 380

DT 01-OCT-2001 (TReMBLrel. 18, Created)
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
 DE HYPOTHETICAL PROTEIN SS02200.
 GN SS02200
 OS Sulfolobus solfataricus;
 CC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 NCBI_TaxID=2287;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=2132296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Ayaaz M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Thierault C., Tojstrup N.,
 RA Charlebois R.L., Doellittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus p2."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL: AE006826; AAK42370.1; -;
 DR InterPro: IPR002789; DUF87.
 DR Pfam: PF01935; DUF87; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 537 AA; 60140 MW; 752F08818FED5D57 CRC64;

Query Match 53.2%; Score 42; DB 17; Length 537;
 Best Local Similarity 43.8%; Pred. No. 1e+02;
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 NRRIRKIIHIKKY 16
 : :|:|:|:|:|
 Db 81 DMNTIKATLHLIKKY 96

RESULT 14
 ID Q9G0I5 PRELIMINARY; PRT; 954 AA.
 AC Q9G0I5;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE ALANYL-TRNA SYNTHETASE.
 GN ALAS.
 OS Giardia lamblia (Giardia intestinalis).
 CC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 CC NCBI_TaxID=5741;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20542069; PubMed=11078517;
 RA Bunjun S., Stachopoulos C., Graham D., Min B., Kitabatake M.,
 RA Wang A.L., Wang C.C., Vivas C.P., Weiss L.M., Soli D.;
 RT "A dual-specificity aminoacyl-tRNA synthetase in the deep-rooted
 RT eukaryote Giardia lamblia."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12997-13002(2000).
 DR EMBL: AF245445; AAG23137.1; -;
 DR InterPro: IPR002318; tRNA-synt_2c.
 DR Pfam: PF01411; tRNA-synt_2c; 2.
 DR PRINTS: PR00980; TRNASYNTHALA.
 KW Aminoacyl-tRNA synthetase.
 SO SEQUENCE 954 AA; 106800 MW; 2F39A8AFBE2C45B8 CRC64;

Query Match 53.2%; Score 42; DB 5; Length 954;
 Best Local Similarity 38.5%; Pred. No. 1.7e+02;
 Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 NRRIRKIIHIKK 13
 : :|:|:|:|:|
 Db 368 NLRNVLRFVPHIL 380

RESULT 15

ID 090022 PRELIMINARY; PRT; 969 AA.
 AC 090022;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ALANYL-TRNA SYNTHETASE (FRAGMENT).
 GN ALAS.
 OS Giardia lamblia (Giardia intestinalis).
 OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.
 OX NCBI_TaxID=5741;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chitade J.W., Brown J.R., Schimmel P., Ribas de Pouplana L.;
 RT *Detection of an intermediate stage of Mitochondria Genesis.*;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF188719; AAF05594.1; -;
 DR InterPro: IPR002318; tRNA-synt_2c.
 DR Pfam: PF01411; tRNA-synt_2c; 4.
 DR PRINTS: PR00980; TRNASYNTHALA.
 KW Aminoacyl-tRNA synthetase.
 FT NON_TER 969
 SQ SEQUENCE 969 AA; 108438 MW; 1A0BCB1EF3780C80 CRC64;

Query Match 53.2%; Score 42; DB 5; Length 969;
 Best Local Similarity 38.5%; Pred. No. 1.7e+02;
 Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 NRRIRIKIHHI 13
 I::I::I::I::
 DB 383 NLRNVLRRVPHIL 395

Search completed: July 12, 2002, 08:03:46
 Job time: 632 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, '08:00:42 ; Search time 227.32 Seconds
(without alignments)
6.841 Million cell updates/sec

Title: US-09-642-744b-13
Perfect score: 66
Sequence: 1 IRRIRKIIHIKK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A-Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	14	22 AAB70660	Ovine SMAP 29 cath
2	66	100.0	14	22 AAB70661	Ovine SMAP 29 cath
3	66	100.0	16	22 AAB70657	Ovine SMAP 29 cath
4	66	100.0	18	22 AAB70654	Ovine SMAP 29 cath
5	66	100.0	20	22 AAB70653	Ovine SMAP 29 cath
6	64	97.0	14	22 AAB70658	Ovine SMAP 29 cath
7	64	97.0	14	22 AAB70659	Ovine SMAP 29 cath
8	64	97.0	16	22 AAB70656	Ovine SMAP 29 cath
9	64	97.0	18	22 AAB70648	Ovine SMAP 29 cath
10	64	97.0	18	22 AAB70655	Ovine SMAP 29 cath
11	64	97.0	29	22 AAB70675	Ovine SMAP 29 cath

12	62	93.9	13	22 AAB70663	Ovine SMAP 29 cath
13	59	89.4	13	22 AAB70662	Ovine SMAP 29 cath
14	57	86.4	12	22 AAB70664	Ovine SMAP 29 cath
15	41	62.1	613	22 AAU35397	Haemophilus influenzae
16	40	60.6	18	22 AAB70652	Ovine SMAP 29 cath
17	40	60.6	18	22 AAG65539	Peptide sequence u
18	40	60.6	28	22 AAB70674	Ovine SMAP 29 cath
19	40	60.6	29	22 AAB70649	Ovine SMAP 29 cath
20	39	59.1	460	22 ABB67825	Drosophila melanog
21	38	57.6	39	20 AA148561	Human breast tumor
22	38	57.6	149	20 AA110929	Human acid sequenc
23	38	57.6	229	21 AA170509	Human BAG-4 protei
24	38	57.6	457	20 AA130348	A suppressor of de
25	38	57.6	457	21 AA170516	Human BAG-4 protei
26	38	57.6	491	22 AAB95343	Human protein sequ
27	38	57.6	519	22 ABB22716	Novel human diagno
28	38	57.6	580	22 ABB07115	Novel human diagno
29	38	57.6	580	22 ABB24597	Novel human diagno
30	38	57.6	1124	22 ABB02951	Novel human diagno
31	38	57.6	1151	22 ABB30207	Novel human diagno
32	38	57.6	1236	21 AAG38618	Arabidopsis thalia
33	37	56.1	17	17 AA105116	Porcine Somatotrop
34	37	56.1	255	22 AAB79587	Cornebacterium g1
35	37	56.1	510	22 AAG90044	C glutamylum prote
36	37	56.1	510	22 AAB79586	Cornebacterium g1
37	37	56.1	844	19 AA143073	HIV-1 gp120 protei
38	36	54.5	27	17 AAR92436	Lytic peptide used
39	36	54.5	27	17 AAR89993	Synthetic lytic pe
40	36	54.5	100	21 AAG23961	Arabidopsis thalia
41	36	54.5	154	22 AA019366	Human G protein-co
42	36	54.5	241	21 AAG37870	Arabidopsis thalia
43	36	54.5	307	21 AAG37869	Arabidopsis thalia
44	36	54.5	314	21 AAG37868	Arabidopsis thalia
45	36	54.5	341	20 AA137104	Protein involved i

ALIGNMENTS

RESULT 1

ID AAB70660 standard; peptide: 14 AA.

AC AAB70660;

DT 15-MAY-2001 (first entry)

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:13.

XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;

KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;

KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;

KW Burkholderia cepacia; Alcaligenes; Xanthomonas.

XX Ovis aries.

XX WO200112668-A1.

PD 22-FEB-2001.

PF 18-AUG-2000: 2000WO-US22781.

PR 18-AUG-1999: 99US-0149886.

PA (IOWA) UNITV IOWA RES FOUND.

PA (REGC) UNITV CALIFORNIA.

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

WP1: 2001-234911/24.

PT New antimicrobial peptides useful as antibiotics for inhibiting growth and proliferation of microbes, and for treating microbial infections -

XX Claim 1; Page 103; 137pp; English.
PS
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
CC
SQ Sequence 14 AA:

Query Match 100.0%; Score 66; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
Db 1 IRRIRKIIHIKK 14

RESULT 2
AAB70661
ID AAB70661 standard; peptide: 14 AA.
XX
AC AAB70661;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:14.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
OS
XX
PN WO200112668-A1.
XX
XX 22-FEB-2001.
PD
XX 18-AUG-2000; 2000WO-US22781.
PF
XX 18-AUG-1999; 99US-0149886.
PR
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI
XX WPI: 2001-234911/24.
DR
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such

CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
CC
SQ Sequence 14 AA:

Query Match 100.0%; Score 66; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
Db 1 IRRIRKIIHIKK 14

RESULT 3
AAB70657
ID AAB70657 standard; peptide: 16 AA.
XX
AC AAB70657;
XX
DT 15-MAY-2001 (first entry)
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:10.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
OS
XX
PN WO200112668-A1.
XX
XX 22-FEB-2001.
PD
XX 18-AUG-2000; 2000WO-US22781.
PF
XX 18-AUG-1999; 99US-0149886.
PR
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI
XX WPI: 2001-234911/24.
DR
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
CC
SQ Sequence 16 AA:

Query Match 100.0%; Score 66; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
|||||

Db 2 irriirkihiik 15

RESULT 4
AAB70654
ID AAB70654 standard; peptide; 18 AA.
XX
XX AAB70654;
AC
XX
XX 15-MAY-2001 (first entry)
XX
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:7.
XX
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
OS
XX WO200112668-A1.
PN
XX 22-FEB-2001.
PD
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
PR
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI; 2001-234911/24.
DR
XX
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
CC
XX
XX Sequence 18 AA:
SQ

Query Match 100.0%; Score 66; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRRIRKIHIIRK 14
| | | | | | | | | | | | | | | | | |
3 IRRIRKIHIIRK 16

Db 3 IRRIRKIHIIRK 16

RESULT 5
AAB70653
ID AAB70653 standard; peptide; 20 AA.
XX
XX AAB70653;
AC
XX
XX 15-MAY-2001 (first entry)
XX
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:6.
XX

XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
OS
XX WO200112668-A1.
PN
XX 22-FEB-2001.
PD
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
PR
XX (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI; 2001-234911/24.
DR
XX
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
CC
XX
XX Sequence 20 AA:
SQ

Query Match 100.0%; Score 66; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRRIRKIHIIRK 14
| | | | | | | | | | | | | | | | | |
3 IRRIRKIHIIRK 16

Db 3 IRRIRKIHIIRK 16

RESULT 6
AAB70658
ID AAB70658 standard; peptide; 14 AA.
XX
XX AAB70658;
AC
XX
XX 15-MAY-2001 (first entry)
XX
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:11.
XX
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
OS
XX WO200112668-A1.
PN
XX 22-FEB-2001.
PD
XX

PT	and proliferation of microbes, and for treating microbial infections -
XX	
PS	Claim 1; Page 103; 137pp; English.
XX	
CC	AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC	AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC	SMP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC	derived from the lupine RCP-18 cathelicidin family peptide. (I) have
CC	antibiotic, antimicrobial and antiviral activities, and can be used as
CC	microbial growth and proliferation inhibitors and in gene therapy. (II)
CC	are useful for inhibiting microbial growth in an environment capable of
CC	sustaining such growth, for inhibiting microbial growth or strain in a
CC	host, and inhibiting the growth of drug-resistant microbial strains such
CC	as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC	Xanthomonas.
XX	
SC	Sequence 14 AA:
OY	Query Match 97.0%: Score 64; DB 22; Length 14;
	Best Local Similarity 92.9%: Pred. No. 0.0007;
	Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0.
OY	1 IRRIRKTIHIKK 14
	:
	Db 1 IRIIRKIHIKK 14
RESULT	8

XX AAB70656;
AC
XX
DT 15-MAY-2001 (first entry)
DE
XX
XX Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:9.
XX
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
bacterial; antibiotic; antiviral; microbial growth inhibitor;
proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
OS
XX
XX WO200112668-A1.
PN
XX
PD 22-FEB-2001.
XX
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI
DR WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
and proliferation of microbes, and for treating microbial infections -
PT
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibacterial, antimicrobial and antiviral activities, and can be used as
CC antimicrobial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a

CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.
 XX
 SQ Sequence 16 AA:

Query Match 97.0%; Score 64; DB 22; Length 16;
 Best Local Similarity 92.9%; Pred. No. 0.00089;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
 :|||||
 DB 2 IRRIRKIIHIKK 15

RESULT 9
 AAB70648
 ID AAB70648 standard; peptide; 18 AA.

AC AAB70648;
 XX
 DT 15-MAY-2001 (first entry)

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:1.

XX Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
 KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
 KM Burkholderia cepacia; Alcaligenes; Xanthomonas.

OS Ovis aries.

XX WO200112668-A1.

PN 22-FEB-2001.

XX 18-AUG-2000; 2000WO-US22781.

PF 18-AUG-1999; 99US-0149886.

PR (IOWA) UNIV IOWA RES FOUND.
 (REGC) UNIV CALIFORNIA.

PA Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

PI WPI, 2001-234911/24.

XX New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -

PS Claim 1; Page 103; 137pp; English.
 XX

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.
 XX

SQ Sequence 18 AA:

Query Match 97.0%; Score 64; DB 22; Length 18;
 Best Local Similarity 92.9%; Pred. No. 0.00089;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14

DB :|||||
 3 IRRIRKIIHIKK 16

RESULT 10
 AAB70655
 ID AAB70655 standard; peptide; 18 AA.

AC AAB70655;

DT 15-MAY-2001 (first entry)

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:8.

XX Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
 KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
 KM Burkholderia cepacia; Alcaligenes; Xanthomonas.

OS Ovis aries.

XX WO200112668-A1.

PN 22-FEB-2001.

PF 18-AUG-2000; 2000WO-US22781.

PR 18-AUG-1999; 99US-0149886.

PA (IOWA) UNIV IOWA RES FOUND.
 (REGC) UNIV CALIFORNIA.

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

PN WPI; 2001-234911/24.

XX New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -

PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
 CC Xanthomonas.
 XX

SQ Sequence 18 AA:

Query Match 97.0%; Score 64; DB 22; Length 18;
 Best Local Similarity 92.9%; Pred. No. 0.00089;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
 :|||||
 DB 3 IRRIRKIIHIKK 16

RESULT 11

AAB70675
 ID AAB70675 standard; peptide; 29 AA.

AC AAB70675;

DT 15-MAY-2001 (first entry)

XX

```

DE  Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:28.
XX
XX  Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KW  bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW  proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW  Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS  Ovis aries.
XX
XX  WO200112668-A1.
XX
XX  22-FEB-2001.
XX
XX  18-AUG-2000; 2000WO-US22781.
XX
XX  18-AUG-1999; 99US-0149886.
XX
XX  (IOWA ) UNIV IOWA RES FOUND.
XX  (REGC ) UNIV CALIFORNIA.
XX
XX  Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX  WPI; 2001-234911/24.
XX
XX  New antimicrobial peptides useful as antibiotics for inhibiting growth
XX  and proliferation of microbes, and for treating microbial infections -
XX  Claim 1; Page 103; 137pp; English.
XX
XX  AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
XX  AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
XX  SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
XX  derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
XX  antibiotic, antimicrobial and antiviral activities, and can be used as
XX  microbial growth and proliferation inhibitors and in gene therapy. (I)
XX  are useful for inhibiting microbial growth in an environment capable of
XX  sustaining such growth, for inhibiting microbial growth or strain in a
XX  host, and inhibiting the growth of drug-resistant microbial strains such
XX  as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX  Xanthomonas.
XX
XX  Sequence 29 AA:
XX
XX  Query Match 97.0%; Score 64; DB 22; Length 29;
XX  Best Local Similarity 92.9%; Pred. No. 0.0014;
XX  Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX  1 IRRIRKRIHIKK 14
XX  :|||||
XX  3 IRIIRKRIHIKK 16
XX
XX  RESULT 12
XX  ID AAB70663
XX  AAB70663 standard; peptide: 13 AA.
XX
XX  AAB70663;
XX
XX  15-MAY-2001 (first entry)
XX
XX  Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:16.
XX
XX  Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KW  bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW  proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW  Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX  Ovis aries.
XX
XX  WO200112668-A1.
XX
XX  22-FEB-2001.

```

```

XX
XX  18-AUG-2000; 2000WO-US22781.
XX
XX  18-AUG-1999; 99US-0149886.
XX
XX  (IOWA ) UNIV IOWA RES FOUND.
XX  (REGC ) UNIV CALIFORNIA.
XX
XX  Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX  WPI; 2001-234911/24.
XX
XX  New antimicrobial peptides useful as antibiotics for inhibiting growth
XX  and proliferation of microbes, and for treating microbial infections -
XX  Claim 1; Page 103; 137pp; English.
XX
XX  AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
XX  AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
XX  SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
XX  derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
XX  antibiotic, antimicrobial and antiviral activities, and can be used as
XX  microbial growth and proliferation inhibitors and in gene therapy. (I)
XX  are useful for inhibiting microbial growth in an environment capable of
XX  sustaining such growth, for inhibiting microbial growth or strain in a
XX  host, and inhibiting the growth of drug-resistant microbial strains such
XX  as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX  Xanthomonas.
XX
XX  Sequence 13 AA:
XX
XX  Query Match 93.9%; Score 62; DB 22; Length 13;
XX  Best Local Similarity 100.0%; Pred. No. 0.0014;
XX  Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX  2 RRIIRKRIHIKK 14
XX  :|||||
XX  1 RRIIRKRIHIKK 13
XX
XX  RESULT 13
XX  ID AAB70662
XX  AAB70662 standard; peptide: 13 AA.
XX
XX  AAB70662;
XX
XX  15-MAY-2001 (first entry)
XX
XX  Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:15.
XX
XX  Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KW  bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW  proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW  Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX  Ovis aries.
XX
XX  WO200112668-A1.
XX
XX  22-FEB-2001.
XX
XX  18-AUG-2000; 2000WO-US22781.
XX
XX  18-AUG-1999; 99US-0149886.
XX
XX  (IOWA ) UNIV IOWA RES FOUND.
XX  (REGC ) UNIV CALIFORNIA.
XX
XX  Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX  WPI; 2001-234911/24.
XX

```

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 XX
 PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (1)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.

XX Sequence 13 AA:

Query Match 89.4%; Score 59; DB 22; Length 13;
 Best Local Similarity 92.3%; Pred. NO. 0.004;
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRIIRKIIHIK 13
 :|||||
 Db 1 RRIIRKIIHIK 13

RESULT 14

AAB70664
 ID AAB70664 standard; peptide: 12 AA.

AC AAB70664;

DT 15-MAY-2001 (first entry)

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:17.

KW Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
 KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
 KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

XX Ovis aries.

OS WO200112668-A1.

PN 22-FEB-2001.

PF 18-AUG-2000; 2000WO-US22781.

XX 18-AUG-1999; 99US-0149886.

XX (IOWA) UNIV IOWA RES FOUND.

PA (REGC) UNIV CALIFORNIA.

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX WPI: 2001-234911/24.

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 XX
 PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (1)
 CC are useful for inhibiting microbial growth in an environment capable of

CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.

XX Sequence 12 AA:

Query Match 86.4%; Score 57; DB 22; Length 12;
 Best Local Similarity 100.0%; Pred. NO. 0.0078;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIK 13
 :|||||
 Db 1 RRIIRKIIHIK 12

RESULT 15

AAU35397
 ID AAU35397 standard; Protein; 613 AA.

AC AAU35397;

DT 14-FEB-2002 (first entry)

DE Haemophilus influenzae cellular proliferation protein #38.

KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.

XX Haemophilus influenzae.

OS WO200170955-A2.

PN 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.

DR N-PDB; AAS53256.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 10990; 51pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an

CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 613 AA;

Query Match 62.1%; Score 41; DB 22; Length 613;
 Best Local Similarity 53.8%; Pred. No. 1e+02;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 2 RRIIRKIIHIKK 14
 ||::|||::||
 Db 353 rllrnehlnkx 365

Search completed: July 12, 2002, 08:00:42
 Job time: 449 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:55:06 ; Search time 75.52 Seconds
(without alignments)
4.528 Million cell updates/sec

Title: US-09-642-744B-13
Perfect score: 66
Sequence: 1 IRRIIRKIHIIKK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCrUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	57.6	457	3	US-09-416-213-2
2	36	54.5	27	2	US-08-505-486-65
3	36	54.5	27	3	US-08-801-028-65
4	36	54.5	27	3	US-09-340-154-65
5	36	54.5	27	5	PCT-US95-09338-65
6	36	54.5	27	5	PCT-US95-09338-65
7	34	51.5	26	1	US-07-725-331-61
8	34	51.5	26	1	PCT-US91-05047-61
9	34	51.5	30	1	US-07-725-331-62
10	34	51.5	30	5	PCT-US91-05047-62
11	34	51.5	36	1	US-07-725-331-63
12	34	51.5	36	5	PCT-US91-05047-63
13	34	51.5	40	1	US-08-179-632-3
14	34	51.5	40	1	US-08-440-174A-3
15	34	51.5	40	5	PCT-US95-00062-3
16	33	50.0	21	1	US-08-786-748A-55
17	33	50.0	21	1	US-08-932-682-55
18	33	50.0	23	2	US-08-505-486-63
19	33	50.0	23	3	US-08-801-028-63
20	33	50.0	23	3	US-09-340-154-63
21	33	50.0	23	5	PCT-US95-09338-63
22	33	50.0	23	5	PCT-US95-09338-63
23	33	50.0	27	2	US-08-505-486-64
24	33	50.0	27	3	US-08-801-028-64
25	33	50.0	27	3	US-09-340-154-64
26	33	50.0	27	5	PCT-US95-09338-64
27	33	50.0	27	5	PCT-US95-09338-64

28	33	50.0	76	2	US-08-117-952-754	Sequence 754, App
29	33	50.0	193	2	US-08-850-119-3	Sequence 3, Appl
30	33	50.0	237	3	US-08-388-353-643	Sequence 643, App
31	33	50.0	237	3	US-08-488-551B-643	Sequence 643, App
32	33	50.0	669	2	US-07-861-800-2	Sequence 2, Appl
33	33	50.0	762	4	US-09-228-986-114	Sequence 114, App
34	33	50.0	855	1	US-08-022-835-6	Sequence 6, Appl
35	33	50.0	855	1	US-08-388-809-6	Sequence 6, Appl
36	33	50.0	855	2	US-08-647-714-6	Sequence 6, Appl
37	33	50.0	1275	3	US-09-120-513-2	Sequence 2, Appl
38	33	50.0	1275	4	US-09-450-105-2	Sequence 2, Appl
39	32.5	49.2	18	4	US-08-992-877-62	Sequence 62, Appl
40	32.5	49.2	475	2	US-08-484-200-2	Sequence 2, Appl
41	32.5	49.2	475	3	US-08-465-375-2	Sequence 2, Appl
42	32.5	49.2	475	4	US-08-764-870-10	Sequence 10, Appl
43	32.5	49.2	475	4	US-08-980-115-10	Sequence 10, Appl
44	32	48.5	18	3	US-09-276-202-7	Sequence 7, Appl
45	32	48.5	294	1	US-08-325-562-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-416-213-2
: Sequence 2, Application US/09416213
: Patent No. 6110690
: GENERAL INFORMATION:
: APPLICANT: Goeddel, David V.
: APPLICANT: Jiang, Yinying
: TITLE OF INVENTION: Suppressors of Death Domains
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
: STREET: 75 DENISE DRIVE
: CITY: HILLSBOROUGH
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94010
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/416,213
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/035,676
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: OSMAN, RICHARD A.
: REGISTRATION NUMBER: 36,627
: TELEPHONE/DOCKET NUMBER: 798-013
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 343-4341
: TELEFAX: (650) 343-4342
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 457 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-416-213-2

Query Match 57.6% Score 38; DB 3; Length 457;
Best Local Similarity 60.0% Pred. No. 78;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 5 IRRIIRKIHIIKK 14
|:|||||:|

Db 380 IKRIIHVLEK 389

RESULT 2

US-08-505-486-65

; Sequence 65, Application US/08505486

; Patent No. 5955573

; GENERAL INFORMATION:

APPLICANT: Jesse M. Jaynes

TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE

TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESS:

ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ

STREET: 555 Thirteenth Street N.W.

CITY: Washington

STATE: D. C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE

COMPUTER: IBM COMPATIBLE

OPERATING SYSTEM: DOS

SOFTWARE: WordPerfect 5.1+

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/505,486

FILING DATE: 21-JUL-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/279,472

FILING DATE: 22-JUL-1994

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: WALKER, BARBARA W.

REGISTRATION NUMBER: 35,400

REFERENCE/DOCKET NUMBER: 2093-117A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)783-6040

TELEFAX: (202)783-6031

INFORMATION FOR SEQ ID NO: 65:

SEQUENCE CHARACTERISTICS:

LENGTH: 27

TYPE: AMINO ACID

TOPOLOGY: LINEAR

MOLECULE TYPE:

DESCRIPTION: PEPTIDE

HYPOTHETICAL: NO

FRAGMENT TYPE: COMPLETE PEPTIDE

ORIGINAL SOURCE: SYNTHETIC

IMMEDIATE SOURCE: SYNTHETIC

PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED

US-08-505-486-65

Query Match 54.5%; Score 36; DB 2; Length 27;

Best Local Similarity 61.5%; Pred. No. 12;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 RRIIRKIIHIKK 14

Db 3 KRIARKIKRIKK 15

RESULT 3

US-08-801-028-65

; Sequence 65, Application US/08801028

; Patent No. 6018102

; GENERAL INFORMATION:

APPLICANT: JOAN GABBARINO

APPLICANT: JESSE M. JAYNES

APPLICANT: WILLIAM BELKAP

TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRODUCTS

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESS:

ADDRESSEE: STEVEN J. HULTQUIST

ADDRESS: INTELLECTUAL PROPERTY/TECHNOLOGY LAW

STREET: 200 PARK DRIVE, SUITE 210

STREET: P.O. BOX 14329

CITY: RESEARCH TRIANGLE PARK

STATE: NORTH CAROLINA

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE

COMPUTER: APPLE MACINTOSH

OPERATING SYSTEM: MACINTOSH

SOFTWARE: M.S. WORD 5.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,028

FILING DATE: 19-FEB-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/279,472

FILING DATE: JULY 22, 1994

APPLICATION NUMBER: 08/225,476

FILING DATE: 04-20-94

APPLICATION NUMBER: 08/225,476

FILING DATE: 04-08-94

APPLICATION NUMBER: 08/039,620

FILING DATE: 06-04-93

APPLICATION NUMBER: 08/148,491

FILING DATE: 11-08-93

APPLICATION NUMBER: 08/148,889

FILING DATE: 11-08-93

ATTORNEY/AGENT INFORMATION:

NAME: WASSERMAN, FRANK S.

REGISTRATION NUMBER: 34,273

REFERENCE/DOCKET NUMBER: 4013-104

TELECOMMUNICATION INFORMATION:

TELEPHONE: (919)990-9531

TELEFAX: (919)990-9532

INFORMATION FOR SEQ ID NO: 65:

SEQUENCE CHARACTERISTICS:

LENGTH: 27

TYPE: AMINO ACID

TOPOLOGY: LINEAR

MOLECULE TYPE:

DESCRIPTION: PEPTIDE

HYPOTHETICAL: NO

FRAGMENT TYPE: COMPLETE PEPTIDE

ORIGINAL SOURCE: SYNTHETIC

IMMEDIATE SOURCE: SYNTHETIC

PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED

US-08-801-028-65

Query Match 54.5%; Score 36; DB 3; Length 27;

Best Local Similarity 61.5%; Pred. No. 12;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 RRIIRKIIHIKK 14

Db 3 KRIARKIKRIKK 15

RESULT 4

US-09-340-154-65

; Sequence 65, Application US/09340154

; Patent No. 6084156

; GENERAL INFORMATION:

APPLICANT: Jesse M. Jaynes

TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE

TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 Thirteenth Street N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: IBM COMPATIBLE
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/340.154
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/505.486
; FILING DATE: 21-JUL-1995
; APPLICATION NUMBER: U.S. 08/279,472
; FILING DATE: 22-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WALKER, BARBARA W
; REGISTRATION NUMBER: 35,400
; REFERENCE/DOCKET NUMBER: 2093-117A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: NO
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
; US-09-340-154-65

Query Match 54.5%; Score 36; DB 3; Length 27;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRIIRKIIHIKK 14
DB 3 KRIARKILKRIKK 15

RESULT 5
PCT-US95-09338-65
; Sequence 65, Application PC/TUS9509338
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; NUMBER OF SEQUENCES: 98
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09338
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA: 08/279,472
; APPLICATION NUMBER: 08/279,472
; FILING DATE: 22-JUL-1994
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
```

```

; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: NO
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
; PCT-US95-09338-65

Query Match 54.5%; Score 36; DB 5; Length 27;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRIIRKIIHIKK 14
DB 3 KRIARKILKRIKK 15

RESULT 7
US-07-725-331-61
; Sequence 61, Application US/07725331
; Patent No. 5294605
; GENERAL INFORMATION:
; APPLICANT: Houghten, Richard

PCT-US95-09339-65
; Sequence 65, Application PC/TUS9509339
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
; NUMBER OF SEQUENCES: 98
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09339
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/279,472
; FILING DATE: 22-JUL-1994
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: NO
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
; PCT-US95-09339-65

Query Match 54.5%; Score 36; DB 5; Length 27;
Best Local Similarity 61.5%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RRIIRKIIHIKK 14
DB 3 KRIARKILKRIKK 15

RESULT 7
US-07-725-331-61
; Sequence 61, Application US/07725331
; Patent No. 5294605
; GENERAL INFORMATION:
; APPLICANT: Houghten, Richard
```



```

? TELEPHONE: 3126165418
? TELEFAX: 3126165460
? INFORMATION FOR SEQ ID NO: 62:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 30 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? FEATURE:
? OTHER INFORMATION: acetylated at N-terminus, may be
? OTHER INFORMATION: a C-terminal amide.
US-07-725-331-62

```

Query Match	51.5%	Score 34:	DB 1:	Length 30:
Best Local Similarity	21.4%	Pred. No. 27:		
Matches	3;	Conservative	10;	Mismatches 1;
				Indels 0;
				Gaps 0

```
Qy 1 IRRIRKIIHIKK 14
    :::::|::: ::||
Db 5 LKKLLKKLLKK 18
```

```

RESULT 10
PCT-US91-05047-62
: Sequence 62: Application PC/TUS9105047
: GENERAL INFORMATION:
: APPLICANT: Houghten, Richard
: APPLICANT: Blondelle, Sylvie
: TITLE OF INVENTION: Amphiphilic Peptide Compositions and
: TITLE OF INVENTION: Analogues Thereof
: NUMBER OF SEQUENCES: 68
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dressler, Goldsmith, Sutter, Shore,
: ADDRESSEE: 6 Milanow
: STREET: 180 North Stetson
: CITY: Chicago
: STATE: IL
: COUNTRY: USA
: ZIP: 60601
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US91/05047
: FILING DATE: 19910717
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/554,422
: FILING DATE: 19-JUL-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Gamson, Edward P.
: REGISTRATION NUMBER: 29,381
: REFERENCE/DOCKET NUMBER: 421250-80
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 3126165418
: TELEFAX: 3126165460
: INFORMATION FOR SEO ID NO: 62:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 30 amino acids
: TYPE: AMINO ACID
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FEATURE:
: OTHER INFORMATION: acetylated at N-terminus, may be
: OTHER INFORMATION: a C-terminal amide.
CCT-US91-05047-62

```

Query Match	51.5%;	Score 34;	DB 5;	Length 30;
Best Local Similarity	21.4%;	Pred. No. 27;		
Matches	3;	Conservative	10;	Mismatches 1;
				Indels 0;
				Gaps 0.
Qy	1	IRRRIRKIIHITKK	14	
	:..:..:..		
Db	5	LKLLKLLKLLKLLK	18	

```

RESULT 11
US-07-725-331-63
; Sequence 63, Application US/07725331
; Patent No. 5294605
; GENERAL INFORMATION:
; APPLICANT: Houghten, Richard
; APPLICANT: Blondelle, Sylvie
; TITLE OF INVENTION: Amphiphilic Peptide Compositions and
; TITLE OF INVENTION: Analogues Thereof
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
; ADDRESSEE: 6 Milanow
; STREET: 180 No. 5294605th Stelson
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/725,331
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/554,422
; FILING DATE: 19-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: 421250-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3126165418
; TELEFAX: 3126165460
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: acetylated at N-terminus, may be
; OTHER INFORMATION: a C-terminal amide
US-07-725-331-63

Query Match 51.5%; Score 34; DB 1; Length 36;
Best Local Similarity 21.4%; Pred. NO. 32;
Matches 3; Conservative 10; Mismatches 1; Indels 0; Gaps 0;

QY 1 IRIIRKIIHIHK 14
; :::::|:::|::|
Db 11 LKRLKRLKLLKK 24

RESULT 12
PCT-US91-05047-63
; Sequence 63, Application PC/TUS9105047
; GENERAL INFORMATION:
; APPLICANT: Houghten, Richard

```

APPLICANT: Blondelle, Sylvie
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
TITLE OF INVENTION: Analogues Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Sulker, Shore,
ADDRESSEE: 6 Milnamow
STREET: 180 North Steetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/05047
FILING DATE: 19910717
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ. ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: AMINO ACID
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
OTHER INFORMATION: a C-terminal amide
PCT-US91-05047-63

Query Match 51.5%; Score 34; DB 5; Length 36;
Best Local Similarity 21.4%; Pred. No. 32;
Matches 3; Conservative 10; Mismatches 1; Indels 0; Gaps 0;

Qy 1 IRRIRKIHIIKK 14
Db 11 LKKLKLKLLK 24

RESULT 13
US-08-179-632-3
Sequence 3, Application US/08179632
Patent No. 5607914
GENERAL INFORMATION:
APPLICANT: Rao, A. Gururaj; Zhong, Lingxiu
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/Microsoft Windows
SOFTWARE: Microsoft Windows No. 5607914epad

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/179,632
FILING DATE: 07-JAN-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/079,512
FILING DATE: 06/18/93
ATTORNEY/AGENT INFORMATION:
NAME: Roth, Michael J.
REGISTRATION NUMBER: 29,342
REFERENCE/DOCKET NUMBER: 0233 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 245-3594
TELEFAX: (515) 245-3634
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-179-632-3

Query Match 51.5%; Score 34; DB 1; Length 40;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 IRRIRKIHIIKK 14
Db 4 VRRIRAIRHIPP 17

RESULT 14
US-08-440-174A-3
Sequence 3, Application US/08440174A
Patent No. 5717061
GENERAL INFORMATION:
APPLICANT: Rao, Gururaj A.
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: 7100 N.W. 62nd Avenue
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,174A
FILING DATE: 12-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/079,512
FILING DATE: 18-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bobrowicz, Donna
REGISTRATION NUMBER: 32,196
REFERENCE/DOCKET NUMBER: 0234R2D-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4896
TELEFAX: (515) 334-6883
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS:

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-174A-3

Query Match 51.5%; Score 34; DB 1; Length 40;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 IRRIRKIIHIKK 14
Db 4 VRRIVRAIRHPRR 17

RESULT 15
PCT-US95-00062-3
Sequence 3, Application PC/TUS9500062
GENERAL INFORMATION:
APPLICANT: Pioneer Hi-Bred International, Inc.
TITLE OF INVENTION: SYNTHETIC ANTIMICROBIAL PEPTIDES
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS/Microsoft Windows
SOFTWARE: Microsoft Windows Notepad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00062
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Yates, Michael E.; Sweeney, Patricia A.;
NAME: Roth, Michael J.; & Simon, Soma G.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 234R2-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00062-3

Query Match 51.5%; Score 34; DB 5; Length 40;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 IRRIRKIIHIKK 14
Db 4 VRRIVRAIRHPRR 17

Search completed: July 12, 2002, 07:55:07
Job time: 114 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:56:52 ; Search time 95.47 Seconds

(without alignments)
14.091 Million cell updates/sec

Title: US-09-642-744B-13

Perfect score: 66

Sequence: 1 IRRIRKIHIIKK 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_71:*
2: pir1:*
3: pir2:*
4: pir3:*
5: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	63.6	125	2 A99114	hypothetical prote
2	42	63.6	867	2 G69485	DNA-directed RNA p
3	41	62.1	337	2 S46010	hypothetical prote
4	41	62.1	613	2 F64056	probable ATP-depen
5	40	60.6	152	2 S68411	cathelin-related p
6	40	60.6	160	2 S68228	myeloid antimicrob
7	40	60.6	160	2 S68412	GTP-binding protei
8	40	60.6	370	2 C70464	hypothetical prote
9	39	59.1	41	2 S77768	conserved hypothet
10	39	59.1	202	1 H70318	hypothetical prote
11	39	59.1	210	2 AC1577	precursor isomer
12	39	59.1	407	2 T37888	hypothetical prote
13	39	59.1	483	2 S82160	hypothetical prote
14	38	57.6	87	2 S80180	sialic protein - Clo
15	38	57.6	234	2 I40822	DNA-dependent RNA
16	38	57.6	234	2 D97108	transposase BME114
17	38	57.6	623	2 S73462	transport ATP-bind
18	38	57.6	698	2 D64084	helicase (EC 3.6.1
19	38	57.6	62	1 RBS3F	ribosomal protein
20	37	56.1	212	2 A64045	KMG-KDPG bifunctio
21	37	56.1	284	2 A10897	tagatose-bisphosph
22	37	56.1	303	2 F71212	hypothetical prote
23	37	56.1	334	2 B72301	endoglucanase - Th
24	37	56.1	364	2 C70322	3-isopropylmalate
25	37	56.1	368	2 C69984	conserved hypothet
26	37	56.1	420	2 S45630	DNA primase chain
27	37	56.1	483	2 S75369	hypothetical prote
28	37	56.1	569	2 C83952	prolyl-tRNA synthe

30	37	56.1	813	2 T20527	hypothetical prote
31	37	56.1	856	1 VCLJ3W	env polyprotein pr
32	36	54.5	67	2 B72248	ribosomal protein
33	36	54.5	102	2 AD1420	PTS celllobiose-spe
34	36	54.5	102	2 AE1795	PTS celllobiose-spe
35	36	54.5	117	2 T12722	hypothetical prote
36	36	54.5	299	2 T23307	hypothetical prote
37	36	54.5	310	2 A33489	hypothetical prote
38	36	54.5	370	2 A05034	hypothetical prote
39	36	54.5	447	2 F81678	GTP-binding protei
40	36	54.5	499	2 G84972	leucyl aminopeptid
41	36	54.5	627	2 C82418	GDEF family prote
42	36	54.5	760	2 T24521	hypothetical prote
43	36	54.5	780	2 AB1801	amino-terminal dom
44	36	54.5	849	2 A96592	hypothetical prote
45	36	54.5	1188	2 C71231	hypothetical prote

ALIGNMENTS

RESULT 1
A99114
hypothetical protein orf125 [imported] - Giardia theta nucleomorph
C:Species: nucleomorph Giardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence-revision 10-May-2001 #text-change 24-May-2001
C:Accession: A99114
R:Donaghy, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.;
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671
A:Accession: A99114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <DOU>
A:Cross-references: GB:AJ010592; NID:g12580762; PIDN:CAC27080.1; GSPDB:GN00151
C:Genetics:
A:Gene: orf125
A:Map position: 2
A:Genome: nucleomorph
C:Keywords: nucleomorph

Query Match 63.6%; Score 42; DB 2; Length 125;
Best Local Similarity 70.0%; Pred. No. 5.9;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 IRRIRKIHIIKK 14
DB 9 IRRIRKIHIIKK 18

RESULT 2
G69485
DNA-directed RNA polymerase, subunit A' (rpoA1) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence-revision 05-Dec-1997 #text-change 18-Jun-1999
C:Accession: G69485
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
glock, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 350, 364-370, 1997
A:Authors: Ullrich, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woes, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343
A:Accession: G69485
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-867 <KLE>
A:Cross-references: GB:AE000972; GB:AE000782; NID:g2689295; PIDN:AAB89365.1; PID:g264
C:Superfamily: Halobacterium DNA-directed RNA polymerase chain A

Query Match 63.6%; Score 42; DB 2; Length 867;
 Best Local Similarity 57.1%; Pred. No. 37;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
 DB 854 VRRIRREIVGEVKK 867

RESULT 3
 S46010
 hypothetical protein YBR14c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YBR118
 C:Species: Saccharomyces cerevisiae
 C:Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 22-Oct-1999
 C:Accession: S46010
 R:Beckam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45995
 A:Accession: S46010
 A:Molecule type: DNA
 A:Residues: 1-337 <BEC>
 A:Cross-references: EMBL:Z36010; NID:g536439; PID:g536440; GSPDB:GM00002; MIPS:YBR141C
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: MIPS:YBR141c
 A:Map position: 2R

Query Match 62.1%; Score 41; DB 2; Length 337;
 Best Local Similarity 61.5%; Pred. No. 22;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14
 DB 30 RRIIRRFHHLINK 42

RESULT 4
 F64056
 Probable ATP-dependent RNA helicase - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb-2001
 C:Accession: F64056
 R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fitchman, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: F64056
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-613 <TIGR>
 A:Cross-references: GB:U32709; GB:I42023; NID:g1573190; PIDN:AAC21900.1; PID:g1573195; T
 C:Keywords: ATP; nucleotide binding; P-loop
 F:49-56/Region: nucleotide-binding motif A (P-loop)
 F:151-156/Region: nucleotide-binding motif B
 F:155-158/Region: DEAD motif

Query Match 62.1%; Score 41; DB 2; Length 613;
 Best Local Similarity 53.8%; Pred. No. 39;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14
 DB 353 RRLRNIEHLMKK 365

RESULT 5
 S68411
 cathelin-related protein 2 precursor - sheep (fragment)
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
 C:Accession: S68411
 R:Mahoney, M.M.; Lee, A.Y.; Brezinski-Caliguri, D.J.; Hutner, K.M.
 FEBS Lett. 377, 519-522, 1995
 A:Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobia
 A:Reference number: S68411; MUID:96140581
 A:Accession: S68411
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-152 <MAH>
 A:Cross-references: EMBL:X92757
 C:Genetics:
 A:Gene: SC5-2
 C:Superfamily: cathelin; cystatin homology
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:14-122/Domain: cystatin homology <CYS>
 F:21-123/Domain: propeptide #status predicted <PRO>
 F:124-152/Product: cathelin-related protein 2 #status predicted <MAT>

Query Match 60.6%; Score 40; DB 2; Length 152;
 Best Local Similarity 57.1%; Pred. No. 15;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
 DB 126 LRRGRKIAHGVRK 139

RESULT 6
 S68228
 myeloid antimicrobial peptide 29 precursor - sheep
 N:Alternate names: cathelicidin
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
 C:Accession: S68228
 R:Bagella, L.; Scocchi, M.; Zanetti, M.
 FEBS Lett. 376, 225-228, 1995
 A:Title: CDNA sequences of three sheep myeloid cathelicidins.
 A:Reference number: S68228; MUID:96105386
 A:Accession: S68228
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-160 <BAG>
 A:Cross-references: EMBL:I46854; NID:g1161248; PIDN:AAA85470.1; PID:g1161249
 C:Superfamily: cathelin; cystatin homology
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:22-130/Domain: cystatin homology <CYS>
 F:29-131/Domain: propeptide #status predicted <PRO>
 F:132-160/Product: myeloid antimicrobial peptide 29 #status predicted <MAT>

Query Match 60.6%; Score 40; DB 2; Length 160;
 Best Local Similarity 57.1%; Pred. No. 16;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
 DB 134 LRRGRKIAHGVRK 147

RESULT 7
 S68412
 cathelin-related protein 1 precursor - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jan-2000
 C:Accession: S68412
 R:Mahoney, M.M.; Lee, A.Y.; Brezinski-Caliguri, D.J.; Hutner, K.M.
 FEBS Lett. 377, 519-522, 1995

A:Title: Molecular analysis of the sheep cathelin family reveals a novel antimicrobial p
A:Reference number: S68411; MUID:96140581
A:Accession: S68412
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-160 <MA>
A:Cross-references: EMBL:X92758
C:Genetics:
A:Gene: SC5-1
C:Superfamily: cathelin; cystatin homology
F:1-29/Domain: signal sequence #status predicted <SIG>
F:22-130/Domain: cystatin homology <Cys>
F:30-131/Domain: propeptide #status predicted <PRO>
F:132-160/Product: cathelin-related protein 1 #status predicted <MAT>

Query Match 60.6% Score 40; DB 2; Length 160;
Best Local Similarity 57.1%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
Db 134 LRLGRKTAHGK 147

RESULT 8
C70464
GTP-binding protein hflX - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Feb-2001
C:Accession: C70464
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666
A:Accession: C70464
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-370 <AQF>
A:Cross-references: GB:AE000762; NID:g2984163; PIDN:AAC07703.1; PID:g2984170; GB:AE00065
A:Experimental source: strain VF5
C:Genetics:
A:Gene: hflX
C:Superfamily: GTP-binding protein hflX; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:197-319/Domain: translation elongation factor Tu homology <ETU>
F:203-210/Region: nucleotide-binding motif A (P-loop)
F:316-319/Region: GTP-binding NKXD motif

Query Match 60.6% Score 40; DB 2; Length 370;
Best Local Similarity 61.5%; Pred. No. 36;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14
Db 159 KRLIKRIRHIKK 171

RESULT 9
S77768
hypothetical protein MC037 - Mycoplasma capricolum (fragment)
C:Species: Mycoplasma capricolum
C:Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 21-Jul-2000
C:Accession: S77768; S48590
R:Bork, P.; Ouzounis, C.; Casati, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.; Mol. Microbiol. 16, 955-967, 1995
A:Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physiology
A:Reference number: S77739; MUID:96059641
A:Accession: S77768
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-41 <BOR>
A:Cross-references: EMBL:Z33033; NID:g541693; PIDN:CAA83711.1; PID:g4379130
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Genetics:
A:Genetic code: SGC3

Query Match 59.1% Score 39; DB 2; Length 41;
Best Local Similarity 50.0%; Pred. No. 6.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
Db 19 LKEFIMQMIHIKK 32

RESULT 10
H70318
conserved hypothetical protein aq_202 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999
C:Accession: H70318
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666
A:Accession: H70318
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-202 <AQF>
A:Cross-references: GB:AE000678; NID:g2982921; PIDN:AAC06551.1; PID:g2982934; GB:AE00
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_202
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein M00226

Query Match 59.1% Score 39; DB 1; Length 202;
Best Local Similarity 45.5%; Pred. No. 30;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 3 RIRKIIHIKK 13
Db 186 RALRKIVHYLK 196

RESULT 11
AC1577
precortin isomerase homolog lin1156 [imported] - Listeria innocua (strain Clp11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AC1577
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Ertlian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maltournam, A.;
Ok, C.; Schueter, T.; Simoes, N.; Tiller, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AC1577
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96387.1; PID:g16413615; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin1156
C:Superfamily: Methanobacterium thermoautotrophicum precortin isomerase

Query Match 59.1%; Score 39; DB 2; Length 210;
 Best Local Similarity 60.0%; Pred. No. 31;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 4 IIRKIIHIKK 13
 ||:||||:|
 Db 62 VIOKIIHVLK 71

RESULT 12

T37888

hypothetical protein SPAC1805.03c - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T37888

R:Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z21752

A:Accession: T37888

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-407 <RFE>

A:Cross-References: EMBL:AL117390; PIDN:CAB55844.1; GSPDB:GN00066; SPDB:SPAC1805.03c

A:Experimental source: strain 972h-; cosmid c1805

C:Genetics:

A:Gene: SPDB:SPAC1805.03c

A:Map position: 1

A:Introns: 32/2; 148/2; 160/3

Query Match

59.1%; Score 39; DB 2; Length 407;
 Best Local Similarity 50.0%; Pred. No. 57;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 RIIRKIIHIKK 14
 |::|||::|

Db 6 RFLRKIIHLVOK 17

RESULT 13

B82160

hypothetical protein VC1762 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)

C:Species: *Vibrio cholerae*

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: B82160

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Winn, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A:Reference number: A82035; MUID:20406833

A:Accession: B82160

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-483 <HEI>

A:Cross-References: GB:AE004253; GB:AE003852; NID:g9656278; PIDN:AAF94911.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1762

A:Map position: 1

Query Match

59.1%; Score 39; DB 2; Length 483;
 Best Local Similarity 53.8%; Pred. No. 67;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14
 ||::||::|

Db 91 RRLNKRITLIVKK 103

RESULT 14

S00180

spermated protein S1 - smaller spotted catshark

C:Species: *Scyliorhinus canicula* (smaller spotted catshark, smaller spotted dogfish)

C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 23-Feb-1997

C:Accession: S00180

R:Chauviere, M.; Martinage, A.; Briand, G.; Sautiere, P.; Chevallier, P.

Eur. J. Biochem. 169, 105-111, 1987

A>Title: Nuclear basic protein transition during sperm differentiation: amino acid se

A:Reference number: S00180; MUID:88055030

A:Accession: S00180

A:Molecule type: protein

A:Residues: 1-87 <CHA>

C:Superfamily: sperm histone

C:Keywords: DNA binding; nucleus; phosphoprotein; sperm

F;3.55/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match

57.6%; Score 38; DB 2; Length 87;
 Best Local Similarity 46.2%; Pred. No. 20;

Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14
 ||::|||::|

Db 42 RRRVKIVHLKRR 54

RESULT 15

I40822

sigK protein - *Clostridium acetobutylicum*

C:Species: *Clostridium acetobutylicum*

C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 21-Jul-2000

C:Accession: I40822

R:Saier, U.; Treuner, A.; Buchholz, M.; Santangelo, J.D.; Durre, P.

J. Bacteriol. 176, 6572-6582, 1994

A>Title: Sporulation and primary sigma factor homologous genes in *Clostridium acetobu*

A:Reference number: I40609; MUID:95050216

A:Accession: I40822

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-234 <RES>

A:Cross-References: GB:I23317; NID:g2745925; PIDN:ANB94775.1; PID:g528974

C:Genetics:

A:Gene: sigK

A:Start codon: GTG

C:Superfamily: transcription initiation factor sigma K; transcription initiation fact

F;56230/Domain: transcription initiation factor sigma katF homology <KTF>

Query Match

57.6%; Score 38; DB 2; Length 234;
 Best Local Similarity 50.0%; Pred. No. 50;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 IIRIRKIIHIKK 14
 |::|::|

Db 55 IERNRLVHLVKK 68

Search completed: July 12, 2002, 07:56:52
 Job time: 219 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:04:38 ; Search time 46.15 seconds
(without alignments)
11.746 Million cell updates/sec

Title: US-09-642-744B-13
Perfect score: 66
Sequence: 1 RRIIRKIHILKK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	62.1	337	1	YB21_YEAST
2	41	62.1	613	1	DEAD_HAEIN
3	40	60.6	160	1	SC51_SHEEP
4	40	60.6	160	1	SC52_SHEEP
5	38	57.6	87	1	SSS1_SCYCA
6	38	57.6	457	1	BAG4_HUMAN
7	38	57.6	623	1	Y014_MYCPN
8	38	57.6	670	1	REP_HAEIN
9	37	56.1	62	1	RL30_BACST
10	37	56.1	212	1	ALKH_HAEIN
11	37	56.1	364	1	LEU3_AQUAE
12	37	56.1	368	1	YSPB_BACSU
13	37	56.1	420	1	PR11_HUMAN
14	37	56.1	847	1	ENV_HV1W2
15	37	56.1	856	1	ENV_HV1W1
16	36	54.5	67	1	RL30_THEMA
17	36	54.5	370	1	MATK_MARO
18	36	54.5	499	1	AMPA_BUCAI
19	36	54.5	590	1	SGT1_SCHPO
20	36	54.5	2210	1	RRPO_LYCAV
21	35.5	53.8	262	1	MURI_BUCAI
22	35	53.0	164	1	RU21_ARARH
23	35	53.0	179	1	CAS2_RAT
24	35	53.0	189	1	TBP_AERPE
25	35	53.0	189	1	TBP_THECE
26	35	53.0	190	1	TBP_PYRKO
27	35	53.0	253	1	RL2_SCHPO
28	35	53.0	417	1	PR11_MOUSE
29	35	53.0	438	1	MCRB_METFE
30	35	53.0	445	1	DHA3_BACSU
31	35	53.0	445	1	PO21_PORJA
32	35	53.0	529	1	TIMK_ECOLI
33	35	53.0	529	1	TIM_SALPO

34	35	53.0	529	1	TIM_SALTY	P40813	salmonella
35	35	53.0	594	1	SYA_BORBU	O51238	borrelia bu
36	35	53.0	622	1	YAG1_SCHPO	O09866	schizosacch
37	35	53.0	672	1	ACSA_PHYBL	O01576	phycomycetes
38	35	53.0	757	1	METE_PASMU	P57843	pasteurella
39	35	53.0	805	1	IF2_AQUAE	O67825	aquifex aeo
40	35	53.0	821	1	MCME_HUMAN	O14566	homo sapien
41	35	53.0	863	1	ENV_HV128	P05882	human immun
42	35	53.0	1229	1	CY43_TRYBB	O99280	trypanosoma
43	35	51.5	90	1	RS20_RICPR	O92CCL	ricketsia
44	34	51.5	130	1	CM36_YEAST	P25603	saccharomyc
45	34	51.5	154	1	YQAR_BACSU	P45914	bacillus su

ALIGNMENTS

RESULT 1
ID YB21_YEAST STANDARD: PRT: 337 AA.
AC P38278;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 38.5 kDa protein in IRA1-MAK5 intergenic region.
GN YBR141C OR YBR1118.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=95274325; PubMed-7754712;
RA Zagulski M., Becam A.-M., Grzybowska E., Lacroix F., Migdalski A.,
RA Slonimski P.P., Sokolowska B., Herbert C.J.;
RT "The sequence of 12.5 kb from the right arm of chromosome II predicts
RT a new N-terminal sequence for the IRA1 protein and reveals two new
RT genes, one of which is a DEAD-box helicase";
RL Yeast 10:1227-1234(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL: Z36010; CAA85099.1; -
CC DR EMBL: X78937; CAA55538.1; -
CC DR PIR: S46010; S46010.
CC DR SGD: S0000345; YBR141C.
CC KW Hypothetical protein.
CC SQ SEQUENCE 337 AA; 38539 MW; CB0039B18FAFB3BE CRC64;

Query Match 62.1%; Score 41; DB 1; Length 337;
Best Local Similarity 61.5%; Pred. No. 4.8;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 RRIIRKIHILKK 14
Db 30 RRIIRFHILKK 42

RESULT 2
ID DEAD_HAEIN STANDARD: PRT: 613 AA.
AC P44586;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Cold-shock DEAD-box protein A homolog (ATP-dependent RNA helicase dead
DE homolog).

GN DEAD OR CSDA OR HI0231.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RD / KW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McEwen K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geophagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";

RL Science 269:496-512(1995).

CC -1- FUNCTION: HAS A HELIX-DESTABILIZING ACTIVITY (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U32709; AAC21900.1; -

DR HSSP: O58083; IHV8.

DR TIGR: H10231; -

DR InterPro: IPR001410; DEAD.

DR InterPro: IPR000629; DEAD_ATP_helicase.

DR InterPro: IPR001650; Helicase_C.

DR Pfam: PF00270; DEAD; 1.

DR Pfam: PF00271; helicase_C; 1.

DR SMART: SM00487; DEXDC; 1.

DR SMART: SM00490; HELIC_C; 1.

DR PROSITE: PS00039; DEAD_ATP_HELICASE; 1.

DR Hydrolase: Helicase; ATP-binding; RNA-binding;
KM Transcription regulation; Complete proteome.

FT NP_BIND 49 56 ATP (POTENTIAL).

FT SITE 155 158 DEAD BOX.

FT SEQUENCE 613 AA; 69705 MW; 1B826CBDEB1704DF CRC64;

SO

Query Match 62.1%; Score 41; DB 1; Length 613;
Best Local Similarity 53.8%; Pred. No. 8.8;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14
II::II::II::II

DB 353 RRLANIEHMKK 365

RESULT 3
SC51_SHEEP STANDARD: PRT; 160 AA.

AC P49928:

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Cathelin-related peptide SC5 precursor 1 (Antibacterial peptide SMAP-
DE 29) (Myeloid antibacterial peptide SMAP-29).

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBI_TaxID=9940;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=96140581; PubMed=8549789;

RA Mahoney M.M., Lee A.Y., Brezinski-Caliguri D.J., Huttner K.M.;

RT "Molecular analysis of the sheep cathelin family reveals a novel
RT antimicrobial peptide.";

RL FEBS Lett. 377:519-522(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA Huttner K.M., Mahoney M.M.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.

CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X92757; CAA63412.1; -

DR EMBL: U60600; AAB49715.1; -

DR InterPro: IPR001894; Cathelicidin.

DR Pfam: PF00666; Cathelicidins; 1.

DR ProDom: PD001838; Cathelicidin; 1.

DR PROSITE: PS00946; CATHELICIDINS_1; 1.

DR PROSITE: PS00947; CATHELICIDINS_2; 1.

DR Antibiotic; Signal.

FT SIGNAL 1 29

FT PROPEP 30 131 BY SIMILARITY.

FT CHAIN 132 160 CATHELIN-RELATED PEPTIDE SC5.

FT MOD_RES 30 30 PYROLIDONE CARBOXYLIC ACID
(BY SIMILARITY).

FT DISULFID 86 97 BY SIMILARITY.

FT DISULFID 108 125 BY SIMILARITY.

FT SEQUENCE 160 AA; 17786 MW; BD9B3859CA32CC49 CRC64;

SO

Query Match 60.6%; Score 40; DB 1; Length 160;
Best Local Similarity 57.1%; Pred. No. 3.3;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 IRIIRKIIHIKK 14
II::II::II::II

DB 134 LRLGRKIAHGVRK 147

RESULT 4
SC52_SHEEP STANDARD: PRT; 160 AA.

AC P49929:

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Cathelin-related peptide SC5 precursor 2 (Antibacterial peptide SMAP-
DE 29) (Myeloid antibacterial peptide SMAP-29).

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBI_TaxID=9940;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=96140581; PubMed=8549789;

RA Mahoney M.M., Lee A.Y., Brezinski-Caliguri D.J., Huttner K.M.;

RT "Molecular analysis of the sheep cathelin family reveals a novel
 RT antimicrobial peptide.";
 RL FEBS Lett. 377:519-522(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Bone marrow;
 RX MEDLINE-96105386; PubMed-7498547;
 RA Bagella L., Scocchi M., Zanetti M.;
 RT "cDNA sequences of three sheep myeloid cathelicidins.";
 RL FEBS Lett. 376:225-228(1995).
 CC -1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.
 CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X92758; CAA63413.1; -;
 DR EMBL: L46854; AAA85470.1; -;
 DR InterPro: IPR001894; Cathelicidin.
 DR InterPro: IPR000010; Cystatin.
 DR Pfam: PF00666; Cathelicidins; 1.
 DR Prodom: PD001838; Cathelicidin; 1.
 DR SMART: SM00043; Cy; 1.
 DR PROSITE: PS00946; CATHELICIDINS_1; 1.
 DR PROSITE: PS00947; CATHELICIDINS_2; 1.
 KW Antidiabetic; Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT PROPEP 30 131 BY SIMILARITY.
 FT CHAIN 132 160 CATHELIN-RELATED PEPTIDE SCS.
 FT MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID
 (BY SIMILARITY).
 FT DISULFID 86 97 BY SIMILARITY.
 FT DISULFID 108 125 BY SIMILARITY.
 FT CONFLICT 28 28 S -> R (IN REF. 2).
 SO SEQUENCE 160 AA; 17742 MW; 4FB98A09355B51F CRC64;

Query Match 60.6%; Score 40; DB 1; Length 160;
 Best Local Similarity 57.1%; Pred. No. 3.3;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 RRIIRKIIHIKK 14
 DB 134 LRRIGRKIIAGVKK 147
 :||: ||| | :||
 SSSL_SCYCA STANDARD; PRT; 87 AA.
 ID SSSL_SCYCA STANDARD; PRT; 87 AA.
 AC P13275;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-JAN-1990 (Rel. 13, Last annotation update)
 DE Spermatid-specific protein SL.
 OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 CC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
 CC Scyllorhinidae; Scyllorhinus.
 CC NCBI_Taxid=7830;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE-88055030; PubMed-3678229;
 RA Chauviere M., Marinage A., Briand G., Sautiere P.,
 RT "Nuclear basic protein transition during sperm differentiation. Amino
 RT acid sequence of a spermatid-specific protein from the dog-fish
 RT Scyllorhinus caniculus.";
 RL Eur. J. Biochem. 169:105-111(1987).
 CC -1- FUNCTION: INVOLVED IN NUCLEAR BASIC PROTEIN TRANSITION: HISTONES

CC ARE REPLACED BY SPERMATID SPECIFIC PROTEINS WHICH ARE THEMSELVES
 CC REPLACED BY PROTAMINES IN LATE SPERMATIDS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: WITH PROTEIN 52.
 DR PIR: S00180; S00180.
 KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
 SW Nuclear protein.
 SO SEQUENCE 87 AA; 11179 MW; 38A0ED7A82536910 CRC64;

Query Match 57.6%; Score 38; DB 1; Length 87;
 Best Local Similarity 46.2%; Pred. No. 4;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIKK 14
 DB 42 RRRVKIYHLKRR 54
 :||: ||| | :||
 RESULT 6
 BAG4_HUMAN
 ID BAG4_HUMAN STANDARD; PRT; 457 AA.
 AC O93429; O95818;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE BAG-family molecular chaperone regulator-4 (Silencer of death
 DE domains).
 GN BAG4 OR SODD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99091615; PubMed-9873016;
 RA Takayama S., Xie Z., Reed J.C.;
 RT "An evolutionarily conserved family of Hsp70/Hsc70 molecular
 RT chaperone regulators.";
 RL J. Biol. Chem. 274:781-786(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99115917; PubMed-9915703;
 RA Jiang Y., Moronicz J.D., Liu W., Goeddel D.V.;
 RT "Prevention of constitutive TNF receptor 1 signaling by silencer of
 RT death domains.";
 RL Science 283:543-546(1999).
 RN [3]
 RP ERRATUM.
 RA Jiang Y., Moronicz J.D., Liu W., Goeddel D.V.;
 RL Science 283:1852-1852(1999).
 CC -1- FUNCTION: INHIBITS THE CHAPERONE ACTIVITY OF HSP70/HSC70 BY
 CC PROMOTING SUBSTRATE RELEASE.
 CC -1- SUBUNIT: BINDS TO THE ATPASE DOMAIN OF HSP70/HSC CHAPERONES.
 CC -1- SIMILARITY: CONTAINS 1 BAG DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF095194; AAD16123.2; -;
 DR EMBL: AF111116; AAD05226.1; -;
 DR MIM: 603884; -;
 DR InterPro: IPR003103; BAG.
 DR Pfam: PF02179; BAG; 1.
 DR SMART: SM00264; BAG; 1.
 KW Chaperone.
 FT DOMAIN 379 456 BAG.
 SO SEQUENCE 457 AA; 49594 MW; BB9D59E8118684A3 CRC64;

Query Match 57.6%; Score 38; DB 1; Length 457;
 Best Local Similarity 60.0%; Pred. No. 22;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 5 IRKIHIKK 14
 |:|||||:
 Db 380 IKIHLVLEK 389

RESULT 7
 Y014_MYCPN STANDARD: PRT: 623 AA.

AC P75095;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical ABC transporter ATP-binding protein MG014 homolog
 DE (D12_ORF623).
 GN MPN018 OR MP136.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RA MEDLINE=9710585; PubMed=8948633;
 RA Himmelfreih R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000015; AAB95784.1; -
 DR HSSP: P13569; INED.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR001140; ABC_transporter_tmam.
 DR InterPro: IPR003439; ABC_transportr.
 DR InterPro: IPR001687; ATP_GTP_A.
 DR Pfam: PF00664; ABC_membrane; 1.
 DR Pfam: PF00005; ABC_tran; 1.
 DR SMART: SM00382; AAA; 1.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
 DR Hypothetical protein; ATP-binding; Transport; Transmembrane;
 KM Complete proteome.
 FT TRANSMEM 27 47 POTENTIAL.
 FT TRANSMEM 86 106 POTENTIAL.
 FT TRANSMEM 157 177 POTENTIAL.
 FT TRANSMEM 180 200 POTENTIAL.
 FT TRANSMEM 266 286 POTENTIAL.
 FT TRANSMEM 307 327 POTENTIAL.
 FT NP_BIND 400 407 ATP (POTENTIAL).
 SO SEQUENCE 623 AA; 68928 MW; A60D4B1D8AEF736B CRC64;

Query Match 57.6%; Score 38; DB 1; Length 623;
 Best Local Similarity 70.0%; Pred. No. 30;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 RRIKRIHI 11
 |:|:|||||:

Db 117 RKIPRIHI 126

RESULT 8
 REP_HAEIN STANDARD: PRT: 670 AA.

AC P44804;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ATP-dependent DNA helicase rep (EC 3.6.1.-).
 GN REP OR H10649.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / RM20 / ATCC 51907;
 RA MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae Rd.";
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: REP HELICASE IS A SINGLE-STRANDED DNA-DEPENDENT
 CC ATPASE INVOLVED IN DNA REPLICATION; IT CAN INITIATE UNWINDING
 CC AT A NICK IN THE DNA. IT BINDS TO THE SINGLE-STRANDED DNA AND
 CC ACTS IN A PROGRESSIVE FASHION ALONG THE DNA IN THE 3' TO 5'
 CC DIRECTION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32748; AAC22309.1; -
 DR HSSP: P09980; UUA.
 DR TIGR: H10649; -
 DR InterPro: IPR000212; UVRD-helicase.
 DR Pfam: PF00580; UVRD-helicase; 1.
 DR HydroLase: Helicase; DNA replication; ATP-binding; DNA-binding;
 KM Complete proteome.
 FT NP_BIND 22 29 ATP (POTENTIAL).
 SO SEQUENCE 670 AA; 77745 MW; 802A131251EDA4E24 CRC64;

Query Match 57.6%; Score 38; DB 1; Length 670;
 Best Local Similarity 61.5%; Pred. No. 32;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 RRIKRIHIKK 14
 |:|:|||||:

Db 30 RVIINKIHLVLEK 42

RESULT 9
 RL30_BACST STANDARD: PRT: 62 AA.

AC P02431;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)

Query Match	56.1%	Score 37	DB 1	Length 62
Best Local Similarity	46.28%	Pred. No. 4.2		
Matches 6	Conservative	4	Mismatches 3	Indels 0
			Gaps	0

RESULT	10
ALKH_HAEIN	
ID	ALKH_HAEIN
STANDARD;	
PRT;	212 AA

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative HMG-KDPG aldolase [includes: 4-hydroxy-2-oxoglutarate
DE aldolase (EC 4.1.3.16) (2-keto-4-hydroxyglutarate aldolase) (HMG-
DE aldolase): 2-dehydro-3-deoxyphosphoglucuronate aldolase (EC 4.1.2.14)
DE (Phospho)-2-dehydro-3-deoxyglucuronate aldolase (Phospho-2-keto-3-
DE deoxyglucuronate aldolase) (2-keto-3-deoxy-6-phosphoglucuronate aldolase)
DE (KDPG-aldolase)].
GN ED4 OR H10047.
OS Haemophilus influenzae.
OS Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OC NCBI_TaxID=727;
RN [1]
RP
RC
RC SRRIN-RD / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodex A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Flierl D.K., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
TL Science 269:496-512(1995).

```

F1          (BI SIMILARITY):
SQ SEQUENCE 212 AA; 22861 MW; 1D920F34AB4C14E3 CRC64;

```

RESULT	11
LEU3_AQUAE	
ID	LEU3_AQUAE
STANDARD;	
PRT;	364 AA

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-Isopropylmalate dehydrogenase (Beta-IPM dehydrogenase)
DE (IMD) (3-IPM-DH).
GN LEUB OR AQ.244.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_Taxid=63363;
RN [1]
RP
RP SEQUENCE FROM N.A.
RC STRAIN=VF5:
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.:
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus".
RL Nucleotide 392353-358(1998).
CC -1- CATALYTIC ACTIVITY: 3-CARBOXY-2-HYDROXY-4-METHYLPENTANATE +
CC NAD(+) = 3-CARBOXY-4-METHYL-2-OXOPENTANATE + NADH (THE PRODUCT
CC DECARBOXYLATES TO 4-METHYL-2-OXOPENTANATE).
CC -1- PATHWAY: THIRD STEP IN LEUCINE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE
CC DEHYDROGENASES FAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000680; AAC06564.1; -.
DR HSSP: P12010; 2AYO.
DR InterPro: IPR001804; Isodh.
DR Pfam: PF00180; Isodh. 1.
DR PROSITE: PS00470; IDH_IMDH: 1.
DR Oxidoreductase; Leucine biosynthesis; NAD; Complete proteome.
KW SEQUENCE 364 AA; 40713 MW; 56B528C760712FAC CRC64;

Query Match 56.1%; Score 37; DB 1; Length 364;
Best Local Similarity 57.1%; Pred. No. 26;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 IRRIRKIHIIKK 14
   |||::|| 1:1
Db 169 IRRIRKGFIRK 182

RESULT 12
YSFB_YACSU STANDARD; PRT; 368 AA.
AC P94536;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical 42.6 kDa protein in sspl-csta intergenic region.
GN YSFB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97124191; PubMed=8969504;
RA Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K., Sanders J.,
RA Emerson P.T., Harwood C.R.;
RT "The dnaB-pheA (256 degrees-240 degrees) region of the Bacillus
RT subtilis chromosome containing genes responsible for stress responses,
RT the utilization of plant cell walls and primary metabolism.";
RL Microbiology 142:3067-3078(1996).
RL -i SIMILARITY: BELONGS TO THE CDAR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 275208; CAA99600.1; -.
DR EMBL: 299118; CAB14827.1; -.
DR Subtilist; BGI2319; ysfB.
KW Hypothetical protein; Complete proteome.
KW SEQUENCE 368 AA; 42611 MW; 37E1C096AF6EF03C CRC64;

Query Match 56.1%; Score 37; DB 1; Length 368;
Best Local Similarity 66.7%; Pred. No. 26;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 IIRIRKIHII 11
   |::|||::||
Db 214 IIRIRKIHII 222

```

```

RESULT 13
PRIL_HUMAN STANDARD; PRT; 420 AA.
AC P49642;
DT 01-FEB-1996 (Rel. 33; Created)
DT 01-FEB-1996 (Rel. 33; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE DNA primase small subunit (EC 2.7.7.-) (DNA primase 49 kDa subunit)
DE (P49).
GN PRIM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94298818; PubMed=8026492;
RA Stadlbauer F., Brueckner A., Reifuss C., Eckerskorn C.,
RA Lotspeich F., Foerster V., Tseng B.Y., Nashauer H.P.;
RT "DNA replication in vitro by recombinant DNA-polymerase-alpha-
RT primase.";
RL Eur. J. Biochem. 222:781-793(1994).
RN [2]
RP SEQUENCE OF 97-146 FROM N.A.
RX MEDLINE=97422622; PubMed=9268648;
RA Cloutier S., Hamel H., Champagne M., Yotov W.V.;
RT "Mapping of the human DNA primase 1 (PRIM1) to chromosome 12q13.";
RL Genomics 43:398-401(1997).
CC -i FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
CC RNA PRIMERS FOR THE OKAZAKI FRAGMENTS MADE DURING DISCONTINUOUS
CC DNA REPLICATION.
CC -i SUBUNIT: HETERODIMER OF A SMALL SUBUNIT AND A LARGE SUBUNIT.
CC -i SIMILARITY: BELONGS TO THE EUKARYOTIC PRIMASE SMALL SUBUNIT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X74330; CAA52377.1; -.
DR EMBL: U89689; AAC51726.1; -.
DR MIM: 176635; -.
DR InterPro: IPR002755; DNA_primase_S.
DR Pfam: PF01896; DNA_primase_S; 1.
KW transferase; DNA replication; DNA-directed RNA polymerase; Primosome.
FT ACT_SITE 44 44 POTENTIAL.
FT ACT_SITE 109 109 POTENTIAL.
FT ACT_SITE 111 111 POTENTIAL.
FT METAL 121 121 POTENTIAL.
FT METAL 122 122 POTENTIAL.
FT METAL 128 128 POTENTIAL.
FT METAL 131 131 POTENTIAL.
FT SEQUENCE 420 AA; 49902 MW; 9B5AC900E0C3CE8 CRC64;

Query Match 56.1%; Score 37; DB 1; Length 420;
Best Local Similarity 64.3%; Pred. No. 30;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 IRRIRKIHIIKK 14
   |||::|||
Db 208 IRRIRKIHIIKK 221

RESULT 14
ENV_HVIW2 STANDARD; PRT; 847 AA.
ID ENV_HVIW2

```

[illegible]

FT	CARBOHYD	450	450	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	602	602	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	607	607	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	616	616	N-LINKED (GLCNAC . .)	(POTENTIAL)
FT	CARBOHYD	628	628	N-LINKED (GLCNAC . .)	(POTENTIAL)
SQ	SEQUENCE	847 AA;	96466 MW;	CDIE33D73MA5BCAE CRC64;	
OY	1 IRRIRKIHIIKK 14				
Db	824 VORICRAIHIIPR 837				
Query Match		Best Local Similarity 56.1%;	Score 37; DB 1;	Length 847;	
Matches 7; Conservative		Pred. No. 62;	Mismatches 4;	Indels 3;	Gaps 0.
RESULT 15					
ENV_HVIM1					
ID ENV_HVM1	STANDARD:	PRT:	856 AA.		
AC P31872;					
DT 01-JUL-1993 (Rel. 26, Created)					
DT 01-JUL-1993 (Rel. 26, Last sequence update)					
DT 15-JUL-1999 (Rel. 38, Last annotation update)					
DE Envelope polyp[protein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].					
GN ENV.					
OS Human immunodeficiency virus type 1 (WMJ1 isolate) (HIV-1).					
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.					
OX NCBI_TaxID=31678;					
RN [1]					
RP SEQUENCE FROM N.A.					
RX MEDLINE=86218077; Pubmed=2423250;					
RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.: "Identification and characterization of conserved and variable regions in the envelope gene of HTLV-II/LAV, the retrovirus of AIDS." Cell 45:637-648(1986).					
RL -I MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SPONTANEOUSLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.					
CC CC					
DR InterPro: IPRO00328; Env_GP41.					
DR InterPro: IPRO00777; GP120.					
DR Pfam: PF00516; GP120: 1.					
DR Pfam: PF00517; GP41: 1.					
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.					
KM SIGNAL	1	29			
FT CHAIN	30	510			
FT CHAIN	511	856			
FT DISULFID	53	73			
FT DISULFID	118	205			
FT DISULFID	125	196			
FT DISULFID	130	152			
FT DISULFID	218	247			
FT DISULFID	228	239			
FT DISULFID	296	330			
FT DISULFID	376	444			
FT DISULFID	383	417			
FT CARBOHYD	87	87			
FT CARBOHYD	134	134			
FT CARBOHYD	140	140			
FT CARBOHYD	151	151			
FT CARBOHYD	155	155			
FT CARBOHYD	183	183			
FT CARBOHYD	197	197			
FT CARBOHYD	234	234			
FT CARBOHYD	241	241			
FT CARBOHYD	262	262			
FT CARBOHYD	276	276			
FT CARBOHYD	289	289			

```

FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA: 97526 MW: DB68D1E49C404DE9 CRC64;

```

```

Query Match 56.1%; Score 37; DB 1; Length 856;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 IRRIRKIHIIKK 14
   :||| | ||| :
Db 833 VQRICRAIHIPRR 846

```

Search completed: July 12, 2002, 08:04:38
Job time: 650 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:03:47 ; Search time 174.7 Seconds

(Without alignments)
13.863 Million cell updates/sec

Title: US-09-642-744b-13

Perfect score: 66

Sequence: 1 IRRIIRKIHIIKK 14

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriaph:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	63.6	125	10	Q9AVZ1 guillardia
2	42	63.6	414	10	Q947U2 oryza sativ
3	42	63.6	500	3	Q9P6R0
4	42	63.6	867	17	Q28391 archaeglob
5	40	60.6	370	16	Q67743 aquifex aeo
6	39	59.1	41	2	Q48971 mycoplasma
7	39	59.1	202	16	Q66580 aquifex aeo
8	39	59.1	210	16	Q92C15 listeria in
9	39	59.1	217	1	Q977R2 unclutered
10	39	59.1	295	3	Q13635 schizosacch
11	39	59.1	407	3	Q9UTR1 schizosacch
12	39	59.1	415	10	Q9SPU4 nicotiana t
13	39	59.1	460	5	Q9VLO3 drosophila
14	39	59.1	483	16	Q9KRB1 vibrio chol
15	38	57.6	56	15	Q56190 human immun
16	38	57.6	56	15	Q56192 human immun

17	38	57.6	147	11	Q91VT5	Q91VT5 mus musculus
18	38	57.6	234	16	Q59264	Q59264 clostridops
19	38	57.6	349	10	Q9SRT5	Q9SRT5 arabidopsis
20	38	57.6	407	10	Q9STW0	Q9STW0 arabidopsis
21	38	57.6	457	11	Q9CWG2	Q9CWG2 mus musculus
22	38	57.6	491	4	Q9H8X2	Q9H8X2 homo sapien
23	38	57.6	852	15	Q69992	Q69992 human immun
24	38	57.6	856	15	Q9E1S5	Q9E1S5 human immun
25	37	56.1	233	2	Q9L7X2	Q9L7X2 clostridium
26	37	56.1	303	17	Q57736	Q57736 pyrococcus
27	37	56.1	334	16	Q9X0D8	Q9X0D8 thermotoga
28	37	56.1	385	2	P94599	P94599 bacillus th
29	37	56.1	483	17	P95941	P95941 sulfolobus
30	37	56.1	569	16	Q9XK71	Q9XK71 bacillus ha
31	37	56.1	578	5	Q95XS3	Q95XS3 caenorhabdi
32	37	56.1	813	5	Q19131	Q19131 caenorhabdi
33	37	56.1	838	15	Q9DVL4	Q9DVL4 human immun
34	37	56.1	856	15	Q41539	Q41539 human immun
35	37	56.1	861	15	Q91U24	Q91U24 human immun
36	37	56.1	867	15	Q74749	Q74749 human immun
37	37	56.1	903	5	Q95NA7	Q95NA7 bulla gould
38	36	54.5	100	17	Q976E3	Q976E3 sulfolobus
39	36	54.5	102	16	Q926Y2	Q926Y2 listeria in
40	36	54.5	117	9	Q80196	Q80196 methanobact
41	36	54.5	216	11	Q9CUL8	Q9CUL8 mus musculu
42	36	54.5	276	11	Q9D3F3	Q9D3F3 mus musculu
43	36	54.5	299	16	Q9X0A0	Q9X0A0 thermotoga
44	36	54.5	314	11	Q9CWM4	Q9CWM4 mus musculu
45	36	54.5	314	11	Q99N96	Q99N96 mus musculu

ALIGNMENTS

RESULT 1	Q9AVZ1	PRELIMINARY;	PRT;	125 AA.
ID	Q9AVZ1			
AC	Q9AVZ1			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)			
DE	HYPOTHETICAL 15.2 KDA PROTEIN.			
OS	Guillardia theta (Cryptomonas phi).			
OC	Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.			
OX	NCBI_TaxID=55529;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20087226; PubMed=10618395;			
RA	Zauner S., Fraunholz M., Mastl J., Penny S.L., Beaton M.,			
RA	Cavaller-Smith T., Maiter U., Douglas S.;			
RT	"Chloroplast protein and centrosomal genes, a tRNA intron, and odd			
RT	telomeres in an unusually compact eukaryotic genome, the cryptomonad			
RT	nucleomorph.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).			
RP	[2]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=21223349; PubMed=11323671;			
RA	Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.,			
RA	Wu X., Reith M., Cavaller-Smith T., Maiter U.;			
RT	"The highly reduced genome of an enslaved algal nucleus.";			
RT	Nature 410:1091-1096(2001).			
DR	EMBL: AT010592; CAC27080.1; -.			
KW	Hypothetical protein.			
SO	SEQUENCE 125 AA; 15207 MW; AAD538B559397A1 CRC64;			
Query Match	63.6%;	Score 42;	DB 10;	Length 125;
Best Local Similarity	70.0%;	Pred. No. 11;		
Matches	7;	Conservative	3;	Mismatches
			0;	Indels
			0;	Gaps
			0;	
Qy	5 IRRIIRK 14			
	:::			
Db	9 IRRIIRK 18			


```

RESULT 2
ID 094702 PRELIMINARY: PRT: 414 AA.
AC 094702:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE TRANSLATION INITIATION FACTOR 2B BETA SUBUNIT.
GN OSJBA0068A07.14.
OS Oryza sativa (Rice).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae;
OC Eurnatoidae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RA de la Bastide M., Spiegel L., Kirchoff K., Preston R., King L.,
RA Vill M.D., Baker J., Zutavern T., Santos L., Miller B., Kuit K.,
RA Cunnius D.M., Bell M., Bailly V., Shah R., Bahret A., Dike S.,
RA Yang C., O'Shaughnessy A., Palmer L., Dedhia N., McCombie W.R.;
RT "Genomic Sequence for Oryza sativa, Nipponbare strain, clone
RT OSJBA0068A07, from Chromosome 10, complete sequence."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC091734; AAK96757.1; -.
KW Initiation factor.
SQ SEQUENCE 414 AA; 44646 MW; BA54863017C7EF0C CRC64;

Query Match 63.6%; Score 42; DB 10; Length 414;
Best Local Similarity 42.9%; Pred. No. 34;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
DB 73 VGNIVRVVLIHIKE 86

RESULT 3
ID 09P6R0 PRELIMINARY: PRT: 500 AA.
AC 09P6R0:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TRANSCRIPTION FACTOR IIIB 70 KDA SUBUNIT.
GN SPBC13E7.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Wood V., Rajadream M.A., Barrell B.G., Skelton J., Church C.M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL354633; CAB89885.1; -.
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GTP-BINDING PROTEIN HFLX.
GN HFLX OR AO_1908.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Feldman D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus."

```

```

DB 247 RRSVREVYHYVK 258

ID 028391 PRELIMINARY: PRT: 867 AA.
AC 028391:
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE DNA-DIRECTED RNA POLYMERASE, SUBUNIT A' (RPOA1).
GN AF1888.
OS Archaeoglobus fulgidus.
OC Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
DR EMBL: AE000972; AAB89365.1; -.
DR TIGR: AF1888; -.
DR InterPro: IPR000722; RNA_POL_A.
DR Pfam: PF00623; RNA_POL_A.1.
KW Hypothetical protein; DNA-directed RNA polymerase; Complete proteome.
SQ SEQUENCE 867 AA; 97806 MW; 6D8051133210085 CRC64;

Query Match 63.6%; Score 42; DB 17; Length 867;
Best Local Similarity 57.1%; Pred. No. 67;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIKK 14
DB 854 VRRIRREVIGEVKK 867

RESULT 5
ID 067743 PRELIMINARY: PRT: 370 AA.
AC 067743:
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GTP-BINDING PROTEIN HFLX.
GN HFLX OR AO_1908.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Feldman D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus."

```

RL Nature 392:353-358(1998).
 DR EMBL: AE000762; AAC07703.1; -.
 KW Complete proteome.
 SQ SEQUENCE 370 AA; 42100 MW; EBD137F19AB1D215 CRC64;

Query Match
 Best Local Similarity 60.6%; Score 40; DB 16; Length 370;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 2 RIIRKIIHIHK 14
 Db 159 RLIRKRIHIHK 171

RESULT 6
 048971 PRELIMINARY; PRT; 41 AA.
 AC 048971;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SIMILAR TO DCTP DEAMINASES (FRAGMENT).
 OS Mycoplasma capricolum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Entomoplasmataceae.
 OX NCBI_TaxID=2095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 27343(KID);
 RX MEDLINE=96059641; PubMed=7476192;
 RA Bork P., Ouzounis C., Casari G., Schneider R., Sander C., Dolan M.,
 RA Gilbert W., Gillevet P.M.;
 RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals
 its physiology."
 RL Mol. Microbiol. 16:955-967(1995).
 DR EMBL: Z33033; CAA83711.1; -.
 FT NON_TER 1 1
 FT NON_TER 41 41
 SQ SEQUENCE 41 AA; 5021 MW; 8BEB3709DA1EB90B CRC64;

Query Match
 Best Local Similarity 59.1%; Score 39; DB 2; Length 41;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Oy 1 IIRIRKIIHIHK 14
 Db 19 LKEFIMOMIHIHK 32

RESULT 7
 066580 PRELIMINARY; PRT; 202 AA.
 AC 066580;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 23.1 KDA PROTEIN.
 GN A0_202.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjlay M., Huber R.,
 RA Felman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus."
 RT Nature 392:353-358(1998).

DR EMBL: AE000678; AAC06551.1; -.
 DR HSSP: Q57679; 1B78.
 DR InterPro: IPR002637; Hamlp_like.
 DR Pfam: PF01725; Hamlp_like; 1.
 DR ProDom: PD004952; Hamlp_like; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 202 AA; 23115 MW; 6709937433F47385 CRC64;

Query Match
 Best Local Similarity 59.1%; Score 39; DB 16; Length 202;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Oy 3 RIIRKIIHIHK 13
 Db 166 RALRKLHVHK 196

RESULT 8
 092CL5 PRELIMINARY; PRT; 210 AA.
 AC 092CL5;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE LIN156 PROTEIN.
 GN LIN156.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Listeria.
 OX NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / SEROVAR 6A;
 RX PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetoui F., Couve E., de Daruvar A., Deloux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussutget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hain J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunz F., Kurapkak G.,
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari R.,
 RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
 RT "Comparative genomics of Listeria species."
 RL Science 294:849-852(2001).
 DR EMBL: AL596167; CAC96387.1; -.
 DR Listlist; LIN01156; -.
 KW Complete proteome.
 SQ SEQUENCE 210 AA; 23107 MW; 205A7B01966F913A CRC64;

Query Match
 Best Local Similarity 59.1%; Score 39; DB 16; Length 210;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 4 IIRKIIHIHK 13
 Db 62 VIQIKIHLK 71

RESULT 9
 0977R2 PRELIMINARY; PRT; 217 AA.
 AC 0977R2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE NAD(P)-FLAVIN OXIDOREDUCTASE.
 OS uncultured crenarchaeote 487.
 OC Archaea; Crenarchaeota; environmental samples;

OC marine archaeal group 1.
 OX NCBI_TaxID=44557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Beja O., Koonin E.V., Aravind L., Taylor L.T., Seltz H., Stein J.L.,
 RA Bensen D.C., Feldman R.A., Swanson R.V., Delong E.F.;
 RT "Comparative genomic analysis of coexisting archaeal genetic variants
 in an Antarctic marine microbial assemblage."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U40238; AAK66808.1; -
 SQ SEQUENCE 217 AA; 24624 MW; E22C83BD4957E67 CRC64;

Query Match 59.1%; Score 39; DB 1; Length 217;
 Best Local Similarity 35.7%; Pred. No. 59;
 Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
 QY 1 IRRIRKIHIHK 14
 :|:|:|:|:|:|
 Db 202 LRQLKEIHVVK 215

RESULT 10
 ID 013635 PRELIMINARY; PRT: 295 AA.
 AC 013635;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL 32.8KD PROTEIN IN NCE3-HHT2 INTERGENIC REGION.
 GN P1043.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX Schizosaccharomycetes.
 NC NCB1_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972 H-;
 RX MEDLINE=20089027; PubMed=10620777;
 RA Machida M., Yamazaki S., Kunihiro S., Tanaka T., Kushiida N., Jinno K.,
 RA Halkawa Y., Yamazaki J., Yamamoto S., Sekine M., Oguchi A., Nagai Y.,
 RA Sakai M., Aoki K., Ogura K., Kudoh Y., Kikuchi H., Zhang M.O.,
 RA Yanagida M.;
 RT "A 38 kb segment containing the cdc2 gene from the left arm of fission
 RT yeast chromosome II: sequence analysis and characterization of the
 RT genomic DNA and cDNAs encoded on the segment."
 RL Yeast 16:71-80(2000).
 DR EMBL; AB004537; BAA21423.1; -
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR SMART; SM00012; PTPc-DSpc; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE_1; UNKNOWN_1.
 SQ SEQUENCE 295 AA; 33211 MW; AC19489DB5DBB20D CRC64;

Query Match 59.1%; Score 39; DB 3; Length 295;
 Best Local Similarity 53.8%; Pred. No. 79;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 IRRIRKIHIHK 13
 |:|:|:|:|:|
 Db 3 IYNIRKVIHLAK 15

RESULT 11
 ID 09UTH1 PRELIMINARY; PRT: 407 AA.
 AC 09UTH1;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE HYPOTHETICAL 47.1 KDA PROTEIN.

GN SPAC1805.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972H-;
 RA Rieger M., McDougall R.C., Rajandream M.A., Barrell B.G.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL117390; CAB55844.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 407 AA; 47107 MW; C1077AEE50E3EE28 CRC64;

Query Match 59.1%; Score 39; DB 3; Length 407;
 Best Local Similarity 50.0%; Pred. No. 11e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 3 IIRIKTIHIHK 14
 |:|:|:|:|:|
 Db 6 RFLRKIAHLVOK 17

RESULT 12
 ID 09SPU4 PRELIMINARY; PRT: 415 AA.
 AC 09SPU4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE TRANSLATION INITIATION FACTOR 2B BETA SUBUNIT.
 GN NIFB.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCB1_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. XANTII;
 RA Park S.-W., Chon G.-H., Yoon G.-M., Ha H.-J., Liu J.-R., Kao T.-H.,
 RA Pallee H.-S.;
 RT "Nifb, encoding PRK1-interacting protein."
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. XANTII;
 RA Park S.-W., Chon G.-H., Yoon G.-M., Ha H.-J., Liu J.-R., Kao T.-H.,
 RA Pallee H.-S.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. XANTII;
 RA Kim M.-I., Park S.-W.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF137288; AAD52847.2; -
 DR InterPro: IPR000649; IF-2B.
 DR Pfam; PF01008; IF-2B; 1.
 KW Initiation factor.
 SQ SEQUENCE 415 AA; 44603 MW; BBH12219935688E3 CRC64;

Query Match 59.1%; Score 39; DB 10; Length 415;
 Best Local Similarity 35.7%; Pred. No. 1.1e+02;
 Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
 QY 1 IRRIRKIHIHK 14
 |:|:|:|:|:|
 Db 73 VGNIVRVHLIIRE 86

```

RESULT 13
O9VL03 PRELIMINARY: PRT; 460 AA.
AC O9VL03:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG13138 PROTEIN.
GN CG13138.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amandides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Daveport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloeck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshireff A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.-A., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003628; AAF52902.1;
DR Flybase: FBgn0032211; CG13138.
SQ SEQUENCE 460 AA; 53135 MW; C038EC3C799814C3 CRC64;

```

```

Query Match 59.1%; Score 39; DB 5; Length 460;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIHK 13
|.:|:|:|:|:|
Db 139 IRIYEVKVIHIHK 151

RESULT 14
O9KR81 PRELIMINARY: PRT; 483 AA.
ID O9KR81

```

```

AC O9KR81:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN VCI1762.
GN VCI1762.
OS Vibrrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Umayam M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL: AE004253; AAF94911.1;
DR TIGR: VCI1762;
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 483 AA; 55076 MW; 34D933F2A8F1DD3C CRC64;

```

```

Query Match 59.1%; Score 39; DB 16; Length 483;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 RIRIRKIIHIHK 14
|.:|:|:|:|:|
Db 91 RRLNKRTIETIVK 103

RESULT 15
O56190 PRELIMINARY: PRT; 56 AA.
ID O56190
AC O56190:
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98127656; PubMed=9468138;
RA Zhu T., Korber B.T., Nambias A.J., Hooper E., Sharp P.M., Ho D.D.;
RT "An African HIV-1 sequence from 1959 and implications for the origin
RT of the epidemic."
RL Nature 391:594-597(1998).
DR EMBL: AF030661; AAC09200.1;
DR InterPro: IPR000328; Env_GP41.
DR Pfam: PF00517; GP41; 1.
KW Transmembrane.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6362 MW; 34C6407BA0EF8FEF CRC64;

Query Match 57.6%; Score 38; DB 15; Length 56;
Best Local Similarity 42.9%; Pred. No. 25;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 IRRIRKIIHIHK 14
|.:|:|:|:|:|
Db 33 VORVFRALIHPRK 46

```

Fri Jul 12 08:56:01 2002

us-09-642-744b-13.rspt

Page 6

Search completed: July 12, 2002, 08:03:48
Job time: 634 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: July 12, 2002, 08:00:42 ; Search time 227.32 Seconds
(without alignments)
7.329 Million cell updates/sec

Title: US-09-642-744B-15

Perfect score: 76

Sequence: 1 RRIIRKIIHIKKNH 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

747574

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/hold-geneseq/genesqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	85.5	20	22	AA70653
2	62	81.6	13	22	AA70663
3	62	81.6	14	22	AA70658
4	62	81.6	14	22	AA70659
5	62	81.6	14	22	AA70660
6	62	81.6	14	22	AA70661
7	62	81.6	16	22	AA70656
8	62	81.6	16	22	AA70657
9	62	81.6	18	22	AA70648
10	62	81.6	18	22	AA70654
11	62	81.6	18	22	AA70655

12	62	81.6	29	22	AA70675	Ovine SMAP 29 cath
13	57	75.0	12	22	AA70664	Ovine SMAP 29 cath
14	57	75.0	13	22	AA70662	Ovine SMAP 29 cath
15	41	53.9	528	16	AA73006	Aminopeptidase O12
16	41	53.9	613	22	AA035397	Haemophilus influenzae
17	41	53.9	1280	22	AB858269	Drosophila melanog
18	40	52.6	18	22	AA655539	Peptide sequence u
19	40	52.6	224	22	AA674358	Human colon cancer
20	40	52.6	477	22	AB859816	Drosophila melanog
21	39	51.3	341	20	AA737104	Protein involved i
22	38	50.0	18	22	AA70652	Ovine SMAP 29 cath
23	38	50.0	28	22	AA70674	Ovine SMAP 29 cath
24	38	50.0	29	22	AA70649	Ovine SMAP 29 cath
25	38	50.0	93	21	AA655537	Arabidopsis thalia
26	38	50.0	109	21	AA655536	Arabidopsis thalia
27	38	50.0	149	20	AA70929	Amino acid sequenc
28	38	50.0	229	21	AA70509	Human BAG-4 protei
29	38	50.0	457	20	AA730348	A suppressor of de
30	38	50.0	457	21	AA70516	Human BAG-4 protei
31	38	50.0	491	22	AA895343	Human protein sequ
32	38	50.0	519	22	AB622716	Novel human diagno
33	38	50.0	728	21	AA636537	Arabidopsis thalia
34	38	50.0	729	21	AA636536	Arabidopsis thalia
35	38	50.0	868	21	AA636535	Arabidopsis thalia
36	38	50.0	1124	22	AB602951	Novel human diagno
37	38	50.0	1151	22	AB630207	Novel human diagno
38	37.5	49.3	18	19	AA663101	Acetylcholine rece
39	37.5	49.3	18	21	AA815076	Acetylcholine rece
40	37.5	49.3	18	21	AA844774	Amino acid sequenc
41	37	48.7	120	22	AB612411	Novel human diagno
42	37	48.7	137	21	AA657782	Arabidopsis thalia
43	37	48.7	179	21	AA657781	Arabidopsis thalia
44	37	48.7	211	22	AA86765	Putative P. abyssal
45	37	48.7	221	22	AB66456	Novel human diagno

ALIGNMENTS

RESULT 1

AA70653 ID AAB70653 standard; peptide: 20 AA.

AC AAB70653:

DT 15-MAY-2001 (first entry)

XX

DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:6.

XX

KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;

KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;

KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;

KW Burkholderia cepacia; Alcaligenes; Xanthomonas.

XX

OS Ovis aries.

OS

PN WO200112668-A1.

XX

PD 22-FEB-2001.

XX

PF 18-AUG-2000; 2000MO-0522781.

XX

PR 18-AUG-1999; 99US-0149886.

XX

PA (IOWA) UNIV IOWA RES FOUND.

PA (REGC) UNIV CALIFORNIA.

XX

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX

DR WPI; 2001-234911/24.

XX

PT New antimicrobial peptides useful as antibiotics for inhibiting growth and proliferation of microbes, and for treating microbial infections -

XX Claim 1; Page 103; 137gp; English.

PS

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which

CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine

CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are

CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have

CC antibiotic, antimicrobial and antiviral activities, and can be used as

CC microbial growth and proliferation inhibitors and in gene therapy. (I)

CC are useful for inhibiting microbial growth in an environment capable of

CC sustaining such growth, for inhibiting microbial growth or strain in a

CC host, and inhibiting the growth of drug-resistant microbial strains such

CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and

CC Xanthomonas.

CC

XX Sequence 20 AA:

SQ

Query Match 85.5%; Score 65; DB 22; Length 20;

Best Local Similarity 88.2%; Pred. No. 0.00069;

Matches 15; Conservative 0; Mismatches 0; Indels 2; Gaps

QY 1 RRIKKIIHIHK--NH 15
|||||||
Db 4 rriirkihikkyghn 20

RESULT 2
AAB70663
ID AAB70663 standard; peptide; 13 AA.
AC AAB70663;
XX 15-MAY-2001 (first entry)
DT
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:16.
XX
XX
XX Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
OS
XX MO200112668-A1.
PN
XX
PD 22-FEB-2001.
PP
XX 18-AUG-2000; 2000WO-US22781.
PF
XX 18-AUG-1999; 99US-0149886.
PR
XX (IOWA) UNIV IOWA RES FOUND.
PA (RSCC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI WPI; 2001-234911/24.
DR
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137gp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which

CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine

CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are

CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have

CC antibiotic, antimicrobial and antiviral activities, and can be used as

CC microbial growth and proliferation inhibitors and in gene therapy. (I)

CC are useful for inhibiting microbial growth in an environment capable of

CC sustaining such growth, for inhibiting microbial growth or strain in a

CC host, and inhibiting the growth of drug-resistant microbial strains such

CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and

CC Xanthomonas.

CC

XX Sequence 20 AA:

SQ

Query Match 85.5%; Score 65; DB 22; Length 20;

Best Local Similarity 88.2%; Pred. No. 0.00069;

Matches 15; Conservative 0; Mismatches 0; Indels 2; Gaps

QY 1 RRIKKIIHIHK--NH 15
|||||||
Db 4 rriirkihikkyghn 20

RESULT 2
AAB70663
ID AAB70663 standard; peptide; 13 AA.
AC AAB70663;
XX 15-MAY-2001 (first entry)
DT
XX
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:16.
XX
XX
XX Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Ovis aries.
OS
XX MO200112668-A1.
PN
XX
PD 22-FEB-2001.
PP
XX 18-AUG-2000; 2000WO-US22781.
PF
XX 18-AUG-1999; 99US-0149886.
PR
XX (IOWA) UNIV IOWA RES FOUND.
PA (RSCC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI WPI; 2001-234911/24.
DR
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137gp; English.

CC	as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC	Xanthomonas.
XX	
SO	Sequence 13 AA:
QY	1 RRIIRKIIRIIRK 13
Db	1 rriirkihikkk 13
RESULT 3	
AAAB70658	
ID	AAAB70658 standard; peptide: 14 AA.
XX	
AC	AAAB70658;
XX	
DT	15-MAY-2001 (first entry)
DE	
XX	Ovine SNAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:11.
KW	Ovine: SNAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
KM	bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KM	proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KM	Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX	
OS	Ovis aries.
XX	
PN	WO200112668-A1.
XX	
PD	22-FEB-2001.
XX	
PF	18-AUG-2000; 2000WO-US22781.
XX	
PR	18-AUG-1999; 99US-0149886.
XX	
XX	(IOWA) UNIV IOWA RES FOUND.
PA	(REGC) UNIV CALIFORNIA.
XX	
PI	Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX	
DR	WPI; 2001-234911/24.
PT	
PT	New antimicrobial peptides useful as antibiotics for inhibiting growth
XX	and proliferation of microbes, and for treating microbial infections
PS	
XX	Claim 1; Page 103; 137pp; English.
XX	
CC	AAAB70648 to AAAB70675 represent antimicrobial peptides (1), of which
CC	AAAB70648 to AAAB70664, AAAB70674 and AAAB70675 are derived from the ovine
CC	SNAP 29 cathelicidin family peptide, and AAAB70665 to AAAB70673 are
CC	derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC	antibiotic, antimicrobial and antiviral activities, and can be used as
CC	microbial growth and proliferation inhibitors and in gene therapy. (1)
CC	are useful for inhibiting microbial growth in an environment capable of
CC	sustaining such growth, for inhibiting microbial growth or strain in a
CC	host, and inhibiting the growth of drug-resistant microbial strains such
CC	as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC	Xanthomonas.
XX	
SO	Sequence 14 AA:
Query Match	81.6%; Score 62; DB 22; Length 14;
Best Local Similarity	100.0%; Pred. No. 0.0015;
Matches 13; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1 RRIIRKIIRIIRK 13	

```

Db      2 rriirkihiik 14

RESULT  4
AAB70659
ID      AAB70659 standard; peptide: 14 AA.
XX
AC      AAB70659;
XX
DT      15-MAY-2001 (first entry)
XX
DE      Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:12.
XX
XX      Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX      bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX      proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX      Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS      Ovis aries.
XX
PN      WO200112668-A1.
XX
PD      22-FEB-2001.
XX
PF      18-AUG-2000; 2000WO-US22781.
XX
PR      18-AUG-1999; 99US-0149886.
XX
PA      (IOWA ) UNIV IOWA RES FOUND.
XX      (REGC ) UNIV CALIFORNIA.
XX
PI      Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX      WPI; 2001-234911/24.
XX
DR      New antimicrobial peptides useful as antibiotics for inhibiting growth
XX      and proliferation of microbes, and for treating microbial infections -
XX
PS      Claim 1; Page 103; 137pp; English.
XX
XX      AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
XX      AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
XX      SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
XX      derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
XX      antibiotic, antimicrobial and antiviral activities, and can be used as
XX      microbial growth and proliferation inhibitors and in gene therapy. (I)
XX      are useful for inhibiting microbial growth in an environment capable of
XX      sustaining such growth, for inhibiting microbial growth or strain in a
XX      host, and inhibiting the growth of drug-resistant microbial strains such
XX      as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX      Xanthomonas.
XX
SQ      Sequence 14 AA:

Query Match      81.6%; Score 62; DB 22: Length 14;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 RRIIRKIHIIRK 13
XX      |||||
XX      2 rriirkihiik 14

RESULT  5
AAB70660
ID      AAB70660 standard; peptide: 14 AA.
XX
AC      AAB70660;
XX
DT      15-MAY-2001 (first entry)
XX
DE      Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:13.

```

```

XX      Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX      bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX      proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX      Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS      Ovis aries.
XX
PN      WO200112668-A1.
XX
PD      22-FEB-2001.
XX
PF      18-AUG-2000; 2000WO-US22781.
XX
PR      18-AUG-1999; 99US-0149886.
XX
PA      (IOWA ) UNIV IOWA RES FOUND.
XX      (REGC ) UNIV CALIFORNIA.
XX
PI      Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX      WPI; 2001-234911/24.
XX
DR      New antimicrobial peptides useful as antibiotics for inhibiting growth
XX      and proliferation of microbes, and for treating microbial infections -
XX
PS      Claim 1; Page 103; 137pp; English.
XX
XX      AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
XX      AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
XX      SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
XX      derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
XX      antibiotic, antimicrobial and antiviral activities, and can be used as
XX      microbial growth and proliferation inhibitors and in gene therapy. (I)
XX      are useful for inhibiting microbial growth in an environment capable of
XX      sustaining such growth, for inhibiting microbial growth or strain in a
XX      host, and inhibiting the growth of drug-resistant microbial strains such
XX      as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX      Xanthomonas.
XX
SQ      Sequence 14 AA:

Query Match      81.6%; Score 62; DB 22: Length 14;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 RRIIRKIHIIRK 13
XX      |||||
XX      2 rriirkihiik 14

RESULT  6
AAB70661
ID      AAB70661 standard; peptide: 14 AA.
XX
AC      AAB70661;
XX
DT      15-MAY-2001 (first entry)
XX
DE      Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:14.
XX
XX      Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX      bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX      proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX      Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS      Ovis aries.
XX
PN      WO200112668-A1.
XX
PD      22-FEB-2001.

```


PT and proliferation of microbes, and for treating microbial infections -
 XX
 PS
 XX

Claim 1, Page 103; 137pp; English.

AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.
 XX
 XX

SO Sequence 16 AA;

Query Match 81.6%; Score 62; DB 22; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.0017;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Oy 1 RRIIRKIIHIKK 13
 ||| ||| ||| ||| |||
 Db 3 rriirkihhiikk 15

RESULT 8

XX AAB70657;
 AC 15-MAY-2001 (first entry)
 DT
 DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:10.
 XX
 XX Ovine: SMAP29; lupine; RCAP 18: cathelicidin; antimicrobial;
 KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
 XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
 XX
 OS Ovis aries.
 OS
 PN WO200112668-A1.
 PD
 PD 22-FEB-2001.
 XX
 XX 18-AUG-2000; 2000WO-US22781.
 XX
 XX 18-AUG-1999; 99US-0149886.
 XX
 XX (IOWA) UNIV IOWA RES FOUND.
 PA (REGC) UNIV CALIFORNIA.
 PA
 XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
 P1 WPI; 2001-234911/24.
 XX
 DR
 PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections
 XX
 XX Claim 1; Page 103; 137pp; English.
 XX
 XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antimicrobial, anticancer and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a

CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.

XX Sequence 16 AA:

Query Match 81.6%; Score 62; DB 22; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.0017;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRIIRKIIHIKK 13
 |||
 Db 3 rrlrkllhlkk 15

RESULT 9
 AAB70648
 ID AAB70648 standard; peptide; 18 AA.

AC AAB70648;
 XX
 DT 15-MAY-2001 (first entry)

XX Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:1.

XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
 KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
 KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

XX Ovis aries.

XX WO200112668-A1.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-US22781.

XX 18-AUG-1999; 99US-0149886.

XX (IOWA) UNIV IOWA RES FOUND.
 PA (REGC) UNIV CALIFORNIA.

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX WPI; 2001-234911/24.

XX New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX Claim 1; Page 103; 137pp; English.

XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.

XX Sequence 18 AA:

Query Match 81.6%; Score 62; DB 22; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.0019;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRIIRKIIHIKK 13

Db |||
 4 rrlrkllhlkk 16

RESULT 10
 AAB70654
 ID AAB70654 standard; peptide; 18 AA.

AC AAB70654;

XX 15-MAY-2001 (first entry)

XX Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:7.

XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
 KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KW proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
 KW *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.

XX Ovis aries.

XX WO200112668-A1.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-US22781.

XX 18-AUG-1999; 99US-0149886.

XX (IOWA) UNIV IOWA RES FOUND.
 PA (REGC) UNIV CALIFORNIA.

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

XX WPI; 2001-234911/24.

XX New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX Claim 1; Page 103; 137pp; English.

XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.

XX Sequence 18 AA:

Query Match 81.6%; Score 62; DB 22; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.0019;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRIIRKIIHIKK 13
 |||
 Db 4 rrlrkllhlkk 16

RESULT 11

AAB70655
 ID AAB70655 standard; peptide; 18 AA.

AC AAB70655;

DT 15-MAY-2001 (first entry)

XX

```

XX PF 18-AUG-2000; 2000WO-US22781.
XX PR 18-AUG-1999; 99US-0149886.
XX PA (IOWA ) UNIV IOWA RES FOUND.
XX PA (RESC ) UNIV CALIFORNIA.
XX PI Tack BE, McCray P, Welsh M, Travis SM, Lehnert R;
XX DR WPI: 2001-234911/24.
XX PT New antimicrobial peptides useful as antibiotics for inhibiting growth
XX PT and proliferation of microbes, and for treating microbial infections -
XX PS Claim 1; Page 103; 137pp; English.
XX CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX SO Sequence 29 AA:

Query Match 81.6%; Score 62; DB 22; Length 29;
Best Local Similarity 100.0%; Pred. No. 0 003;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RRIRKRIHIIRK 13
| | | | | | | | | |
DB 4 RLIRKIHLIKK 16

RESULT 13
AAB70664
ID AAB70664 standard; peptide: 12 AA.
AC AAB70664;
AD
DT 15-MAY-2001 (first entry)
DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:17.
DX
XX OVine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX OS Ovis aries.
XX PN WO200112668-A1.
XX PD 22-FEB-2001.
XX PF 18-AUG-2000; 2000WO-US22781.
XX PR 18-AUG-1999; 99US-0149886.
XX PA (IOWA ) UNIV IOWA RES FOUND.
XX PA (RESC ) UNIV CALIFORNIA.
XX PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX DR WPI: 2001-234911/24.

```

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 PS Claim 1; Page 103; 137pp; English.
 XX
 CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.
 CC
 SO Sequence 12 AA:
 XX
 Query Match 75.0%; Score 57; DB 22; Length 12;
 Best Local Similarity 100.0%; Pred. No. 0.0079;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RRIIRKIIHIK 12
 Db 1 RRIIRKIIHIK 12
 XX
 RESULT 14
 AAB70662
 ID AAB70662 standard; peptide; 13 AA.
 XX
 AC AAB70662;
 XX
 DT 15-MAY-2001 (first entry)
 XX
 DE Ovine SMAP 29 cathelicidin derived antimicrobial peptide SEQ ID NO:15.
 XX
 KM Ovine: SMAP29; lupine: RCAP 18; cathelicidin; antimicrobial;
 KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
 KM proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
 KM *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
 XX
 OS Ovis aries.
 XX
 PN WO200112668-A1.
 XX
 PD 22-FEB-2001.
 XX
 PF 18-AUG-2000; 2000WO-US22781.
 XX
 PR 18-AUG-1999; 99US-0149886.
 XX
 PA (IOWA) UNIV IOWA RES FOUND.
 PA (RECG) UNIV CALIFORNIA.
 XX
 PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
 XX
 DR WPI; 2001-234911/24.
 XX
 PT New antimicrobial peptides useful as antibiotics for inhibiting growth
 PT and proliferation of microbes, and for treating microbial infections -
 XX
 PS Claim 1; Page 103; 137pp; English.
 XX
 CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of

CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.
 CC
 SO Sequence 13 AA:
 XX
 Query Match 75.0%; Score 57; DB 22; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.0085;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 RRIIRKIIHIK 12
 Db 2 RRIIRKIIHIK 13
 XX
 RESULT 15
 AAR73006
 ID AAR73006 standard; Protein; 528 AA.
 XX
 AC AAR73006;
 XX
 DT 30-OCT-1995 (first entry)
 XX
 DE Aminopeptidase O12 clone O12-4.
 XX
 KW Aminopeptidase; O12; antigen; helminth; nematode; parasite;
 KW vaccine.
 XX
 OS Ostertagia ostertagi.
 OS
 OS WO9512671-A.
 PN
 PD 11-MAY-1995.
 XX
 PF 03-NOV-1994; 94WO-GB02414.
 XX
 PR 03-NOV-1993; 93GB-0022702.
 XX
 PA (BIOT-) BIOTECHNOLOGY & BIOLOGICAL SCI RES COUNC.
 PA
 PI Graham M, McMichael-Phillips D, Munn EA;
 XX
 DR WPI; 1995-193820/25.
 DR N-PSDB; AA086997.
 XX
 PT A helminth parasite antigen with aminopeptidase like activity - for
 PT use in vaccines to stimulate an immune response against helminth
 PT parasites pref. gastric nematodes of the gastro-intestinal tracts of
 PT mammals
 XX
 PS Disclosure; Page 70-72; 127pp; English.
 XX
 CC Primers based on helminth microsomal aminopeptidases were used for
 CC the PCR amplification of O. osteragi cDNA. Sequencing of PCR
 CC products identified 5 closely-related O12 sequences, O12-1 to -5
 CC (given in AA086994-98).
 XX
 SO Sequence 528 AA;
 XX
 Query Match 53.9%; Score 41; DB 16; Length 528;
 Best Local Similarity 72.7%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 5 RKIIRKIIKXN 15
 Db 294 RKIIRKIIKXN 304

Job time: 449 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:55:07 ; Search time 75.52 Seconds
(without alignments)
4.851 Million cell updates/sec

Title: US-09-642-744B-15
Perfect score: 76
Sequence: 1 RRIIRKIIHIKKNH 15

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	38	50.0	457	3	US-09-416-213-2
2	37.5	49.3	18	4	US-08-992-877-62
3	37	48.7	682	1	US-08-441-139-2
4	36	47.4	27	2	US-08-505-486-65
5	36	47.4	27	3	US-08-801-028-65
6	36	47.4	27	3	US-09-340-154-65
7	36	47.4	27	5	PCT-US95-09338-65
8	36	47.4	27	5	PCT-US95-09339-65
9	36	47.4	65	1	US-07-879-685B-1
10	36	47.4	162	1	US-07-879-685B-4
11	35	46.1	104	4	US-09-029-213B-21
12	35	46.1	281	4	US-09-404-258-17
13	35	46.1	483	4	US-09-106-194-4
14	35	46.1	531	2	US-08-975-114A-4
15	35	46.1	531	4	US-08-849-281A-4
16	35	46.1	574	4	US-09-079-431B-6
17	35	46.1	603	4	US-08-687-865A-2
18	35	46.1	629	4	US-09-043-711-2
19	35	46.1	629	4	US-09-079-431B-4
20	35	46.1	630	4	US-09-079-431B-2
21	35	46.1	685	3	US-09-031-563-21
22	35	46.1	1055	3	US-09-031-563-27
23	35	46.1	1315	3	US-09-031-563-22
24	35	46.1	1315	3	US-09-031-563-25
25	35	46.1	1315	4	US-09-293-505-10
26	35	46.1	3177	2	US-08-477-451-4
27	34	44.7	297	4	US-09-068-195-12

28	34	44.7	456	2	US-08-819-013-1	Sequence 1, Appl
29	34	44.7	536	2	US-08-975-114A-5	Sequence 5, Appl
30	34	44.7	536	4	US-08-849-281A-3	Sequence 4, Appl
31	34	44.7	643	2	US-08-551-356-4	Sequence 3, Appl
32	34	44.7	643	5	PCT-US93-12687-4	Sequence 4, Appl
33	34	44.7	644	1	US-08-206-176-2	Sequence 2, Appl
34	34	44.7	661	3	US-09-022-875-2	Sequence 2, Appl
35	34	44.7	1164	4	US-09-457-708-2	Sequence 2, Appl
36	33	43.4	15	1	US-08-193-521-7	Sequence 7, Appl
37	33	43.4	15	1	US-08-434-120-101	Sequence 101, App
38	33	43.4	15	1	US-08-465-325-100	Sequence 100, App
39	33	43.4	15	4	US-09-115-737-100	Sequence 100, App
40	33	43.4	18	2	US-08-497-599-24	Sequence 24, Appl
41	33	43.4	57	4	US-08-630-915A-206	Sequence 206, App
42	33	43.4	76	2	US-08-117-952-754	Sequence 754, App
43	33	43.4	102	3	US-08-478-097A-17	Sequence 17, Appl
44	33	43.4	102	4	US-08-931-858E-162	Sequence 162, App
45	33	43.4	102	4	US-08-981-739-162	Sequence 162, App

ALIGNMENTS

```

RESULT 1
US-09-416-213-2
; Sequence 2, Application US/09416213
; Patent No. 6110690
GENERAL INFORMATION:
APPLICANT: Goeddel, David V.
APPLICANT: Jiang, Yingling
TITLE OF INVENTION: Suppressors of Death Domains
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/416,213
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/035,676
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: T98-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 457 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-416-213-2
Query Match 50.0%; Score 38; DB 3; Length 457;
Best Local Similarity 60.0%; Pred. No. 99;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 4 IRKIIHIKK 13
I:||||:::|

```

DB 380 IKKIHVLEK 389

RESULT 2

US-08-992-877-62

Sequence 62, Application US/08992877

Patent No. 6340461

GENERAL INFORMATION:

APPLICANT: Terman, David S

TITLE OF INVENTION: SUPERANTIGEN BASED METHODS AND COMPOSITIONS FOR TREATMENT OF INFECTIOUS DISEASE

FILE REFERENCE: Superantigen

CURRENT APPLICATION NUMBER: US/08/992,877

CURRENT FILING DATE: 1997-12-17

PRIOR APPLICATION NUMBER: 60/044,074

PRIOR FILING DATE: 1997-04-17

NUMBER OF SEQ ID NOS: 78

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 62

LENGTH: 18

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:antigen

US-08-992-877-62

Query Match 49.3%; Score 37.5; DB 4; Length 18;

Best Local Similarity 61.5%; Pred. No. 5.2;

Matches 8; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

OY 3 IIRKIHIHKKH 15

1::1111:111

DB 6 IVKK-IHIVKKH 17

RESULT 3

US-08-441-139-2

Sequence 2, Application US/08441139

Patent No. 5773245

GENERAL INFORMATION:

APPLICANT: Wittup, Dr. Karl D.

APPLICANT: Robinson, Anne S.

TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF RECOMBINANTLY EXPRESSED PROTEINS

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: NY

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/441,139

FILING DATE: 15-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/089,997

FILING DATE: 06-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Digilio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 8646

TELECOMMUNICATION INFORMATION:

TELEPHONE: 516-742-4343

TELEFAX: 516-742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 682 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-441-139-2

Query Match 48.7%; Score 37; DB 1; Length 682;

Best Local Similarity 42.9%; Pred. No. 2.1e+02;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 IIRKIHIHKKH 15

1::1:111

DB 281 KIVROLIKAFKKH 294

RESULT 4

US-08-505-486-65

Sequence 65, Application US/08505486

Patent No. 5955573

GENERAL INFORMATION:

APPLICANT: Jesse M. Jaynes

TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE

TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND METHODS OF MAKING AND USING SAME

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESS:

ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ

STREET: 555 Thirteenth Street N.W.

CITY: Washington

STATE: D. C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE

COMPUTER: IBM COMPATIBLE

OPERATING SYSTEM: DOS

SOFTWARE: Wordperfect 5.1+

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/505,486

FILING DATE: 21-JUL-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/279,472

FILING DATE: 22-JUL-1994

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: WALKER, BARBARA W.

REGISTRATION NUMBER: 35,400

REFERENCE/DOCKET NUMBER: 2093-117A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)783-6040

TELEFAX: (202)783-6031

INFORMATION FOR SEQ. ID NO: 65:

SEQUENCE CHARACTERISTICS:

LENGTH: 27

TYPE: AMINO ACID

TOPOLOGY: LINEAR

MOLECULE TYPE:

DESCRIPTION: PEPTIDE

HYPOTHETICAL: NO

FRAGMENT TYPE: COMPLETE PEPTIDE

ORIGINAL SOURCE: SYNTHETIC

IMMEDIATE SOURCE: SYNTHETIC

PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED

US-08-505-486-65

Query Match 47.4%; Score 36; DB 2; Length 27;

Best Local Similarity 61.5%; Pred. No. 13;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRIKKIHIKK 13
:|||:|
Db 3 KRIARKILKRIKK 15

RESULT 5
US-08-801-028-65
; Sequence 65, Application US/08801028
; Patent No. 6018102
; GENERAL INFORMATION:
; APPLICANT: JOAN GARBARINO
; APPLICANT: JESSE M. JAYNES
; APPLICANT: WILLIAM BELKNAP
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRODUCT
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
; STREET: 200 PARK DRIVE, SUITE 210
; STREET: P.O. BOX 14329
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,028
; FILING DATE: 19-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/279,472
; FILING DATE: JULY 22, 1994
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-20-94
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-08-94
; APPLICATION NUMBER: 08/039,620
; FILING DATE: 06-04-93
; APPLICATION NUMBER: 08/148,491
; FILING DATE: 11-08-93
; APPLICATION NUMBER: 08/148,889
; FILING DATE: 11-08-93
; ATTORNEY/AGENT INFORMATION:
; NAME: MASSEMAN, FRAN S.
; REGISTRATION NUMBER: 34,273
; REFERENCE/DOCKET NUMBER: 4013-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: NO
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
; US-08-801-028-65

Query Match 47.4%; Score 36; DB 3; Length 27;
Best Local Similarity 61.5%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRIKKIHIKK 13
:|||:|
Db 3 KRIARKILKRIKK 15

RESULT 6
US-09-340-154-65
; Sequence 65, Application US/09340154
; Patent No. 6084156
; GENERAL INFORMATION:
; APPLICANT: Jesse M. Jaynes
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 Thirteenth Street N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: IBM COMPATIBLE
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/340,154
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/505,486
; FILING DATE: 21-JUL-1995
; APPLICATION NUMBER: U.S. 08/279,472
; FILING DATE: 22-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WALKER, BARBARA W.
; REGISTRATION NUMBER: 35,400
; REFERENCE/DOCKET NUMBER: 2093-117A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: NO
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
; US-09-340-154-65

Query Match 47.4%; Score 36; DB 3; Length 27;
Best Local Similarity 61.5%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRIKKIHIKK 13
:|||:|
Db 3 KRIARKILKRIKK 15

RESULT 7
PCT-US95-09338-65
; Sequence 65, Application PC/TUS9509338
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE


```

: TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
: TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
: NUMBER OF SEQUENCES: 98
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORDPERFECT 5.1+
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/09338
: FILING DATE: 21-JUL-1994
: PRIOR APPLICATION DATA: 08/279,472
: APPLICATION NUMBER: 08/279,472
: FILING DATE: 22-JUL-1994
: INFORMATION FOR SEQ ID NO: 65:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 27
: TYPE: AMINO ACID
: TOPOLOGY: LINEAR
: MOLECULE TYPE:
: DESCRIPTION: PEPTIDE
: HYPOTHETICAL: NO
: FRAGMENT TYPE: COMPLETE PEPTIDE
: ORIGINAL SOURCE: SYNTHETIC
: IMMEDIATE SOURCE: SYNTHETIC
: PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US95-09338-65

Query Match          47.4%; Score 36; DB 5; Length 27;
Best Local Similarity 61.5%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRIKKIHIKK 13
   :| | | | | | |
Db 3 KRIARKIIRIKK 15

RESULT 8
PCT-US95-09339-65
: Sequence 65, Application PC/TUS9509339
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
: TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
: NUMBER OF SEQUENCES: 98
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORDPERFECT 5.1+
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/09339
: FILING DATE: 21-JUL-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/279,472
: FILING DATE: 22-JUL-1994
: INFORMATION FOR SEQ ID NO: 65:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 27
: TYPE: AMINO ACID
: TOPOLOGY: LINEAR
: MOLECULE TYPE:
: DESCRIPTION: PEPTIDE
: HYPOTHETICAL: NO
: FRAGMENT TYPE: COMPLETE PEPTIDE
: ORIGINAL SOURCE: SYNTHETIC
: IMMEDIATE SOURCE: SYNTHETIC
: PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US95-09339-65
```

```

Query Match          47.4%; Score 36; DB 5; Length 27;
Best Local Similarity 61.5%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRIKKIHIKK 13
   :| | | | | | |
Db 3 KRIARKIIRIKK 15

RESULT 9
US-07-879-685B-1
: Sequence 1, Application US/07879685B
: Patent No. 5296383
: GENERAL INFORMATION:
: APPLICANT: DAIKIN INDUSTRIES, LTD.
: TITLE OF INVENTION: A human centromere antigen
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Umeda Center Building, 4-12
: STREET: Nakazaki-nishi, 2-chome
: CITY: Kita-ku
: STATE: Osaka
: COUNTRY: Japan
: ZIP: 530
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/879,685B
: FILING DATE: 19920507
: CLASSIFICATION: 436
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 3-102517
: FILING DATE: 08-May-1991
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 65 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FRAGMENT TYPE: Internal fragment
: ORIGINAL SOURCE:
: ORGANISM: human
US-07-879-685B-1

Query Match          47.4%; Score 36; DB 1; Length 65;
Best Local Similarity 44.4%; Pred. No. 31;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 7 IHIKKNH 15
   :| | | | | | |
Db 44 LVHVTIRKNH 52

RESULT 10
US-07-879-685B-4
: Sequence 4, Application US/07879685B
: Patent No. 5296383
: GENERAL INFORMATION:
: APPLICANT: DAIKIN INDUSTRIES, LTD.
: TITLE OF INVENTION: A human centromere antigen
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Umeda Center Building, 4-12
: STREET: Nakazaki-nishi, 2-chome
: CITY: Kita-ku
: STATE: Osaka
```

COUNTRY: Japan
ZIP: 350
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/879,685B
FILING DATE: 19920507
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-102517
FILING DATE: 08-May-1991
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-879-685B-4

Query Match 47.4%; Score 36; DB 1; Length 162;
Best Local Similarity 44.4%; Pred. No. 74;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 7 IIRIKKNN 15
DB 141 LVHTYRKNH 149

RESULT 11
US-09-029-213B-21
Sequence 21, Application US/09029213B
GENERAL INFORMATION:
APPLICANT: CHRISTIAN, Peter D.
TITLE OF INVENTION: RECOMBINANT HELICOVERPA BACULOVIRUSES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDermott, Will & Emery
STREET: 600 13th Street, NW
CITY: Washington
STATE: District of Columbia
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/029,213B
FILING DATE: 31-AUG-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Joseph Hyosuk Kim
REGISTRATION NUMBER: 41,425
REFERENCE/DOCKET NUMBER: 50179-048
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-756-8000
TELEFAX: 202-756-8087
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-029-213B-21

Query Match 46.1%; Score 35; DB 4; Length 104;
Best Local Similarity 58.3%; Pred. No. 69;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 3 IIRIKIIRIKNN 14
DB 59 IYRRIIELEIKNN 70

RESULT 12
US-09-404-258-17
Sequence 17, Application US/09404258
Patent No. 6274353
GENERAL INFORMATION:
APPLICANT: YANG, Shuwei
TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR IMPROVED POLYNUCLEOTIDE
FILE REFERENCE: 089962/0102
CURRENT APPLICATION NUMBER: US/09/404,258
CURRENT FILING DATE: 1999-09-22
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 17
LENGTH: 281
TYPE: PRT
ORGANISM: Pyrococcus abyssi
US-09-404-258-17

Query Match 46.1%; Score 35; DB 4; Length 281;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 IIRIKIIRIKNN 14
DB 52 LARKIRKIVAKNN 63

RESULT 13
US-09-106-194-4
Sequence 4, Application US/09106194
Patent No. 6262234
GENERAL INFORMATION:
APPLICANT: Holloway, James
APPLICANT: Jelinek, Laura
APPLICANT: Blumberg, Hal
TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR POLYPEPTIDE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/106,194
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Leith, Debra K
REGISTRATION NUMBER: 32,619
REFERENCE/DOCKET NUMBER: 96-11

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6674
TELEFAX: 206-442-6678
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-106-194-4

Query Match
Best Local Similarity 46.1%; Score 35; DB 4; Length 483;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 9 HIKKNH 15
|:|:|:|
Db 418 HVIQKNH 424

RESULT 14
US-08-975-114A-4
Sequence 4, Application US/08975114A
Patent No. 5876714
GENERAL INFORMATION:
APPLICANT: Acsushi NISHIKAWA et al.
TITLE OF INVENTION: HUMAN GLYCOSYLTRANSFERASE GENE,
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,114A
FILING DATE: No. 5876714ember 20, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/107,173
FILING DATE: August 17, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 1-P3439DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 531 amino acid residues
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE:
US-08-975-114A-4

Query Match
Best Local Similarity 46.1%; Score 35; DB 2; Length 531;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRIIRKIIHIKKNH 15
|:|:|:|:|
Db 200 REVPRRVINAINVNH 214

RESULT 15
US-08-849-281A-4
Sequence 4, Application US/08849281A
Patent No. 6153433
GENERAL INFORMATION:
APPLICANT: Eiji MIYOSHI et al.
TITLE OF INVENTION: INHIBITOR FOR VIRAL REPLICATION
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
STREET: 2033 K Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,281A
FILING DATE: May 30, 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lee Cheng
REGISTRATION NUMBER: 40,949
REFERENCE/DOCKET NUMBER: 97-0529*/LC(MJ)/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-721-8200
TELEFAX: 202-721-8250
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 531 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
US-08-849-281A-4

Query Match
Best Local Similarity 46.1%; Score 35; DB 4; Length 531;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 RRIIRKIIHIKKNH 15
|:|:|:|:|
Db 200 REVPRRVINAINVNH 214

Search completed: July 12, 2002, 07:55:08
Job time: 115 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:56:52 ; Search time 95.47 Seconds

(Without alignments)
15.097 Million cell updates/sec

Title: US-09-642-744B-15

Perfect score: 76

Sequence: 1 RRIIRKIIHIKKNH 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	57.9	623	2	S73462
2	42	55.3	125	2	A99114
3	42	55.3	310	2	A33489
4	41	53.9	95	2	T21413
5	41	53.9	305	2	A10847
6	41	53.9	337	2	S46010
7	41	53.9	613	2	F64056
8	41	53.9	1286	2	T18734
9	40	52.6	70	2	G90478
10	40	52.6	370	2	C70464
11	40	52.6	455	2	T06357
12	39	51.3	87	2	S00180
13	39	51.3	202	1	H70318
14	39	51.3	203	2	H89768
15	39	51.3	206	2	F64620
16	39	51.3	206	2	E71894
17	39	51.3	210	2	AC1577
18	39	51.3	310	2	AG1733
19	39	51.3	334	2	B72301
20	39	51.3	407	2	T37888
21	39	51.3	431	1	EB7BDP
22	39	51.3	440	2	F99708
23	39	51.3	483	2	B82160
24	39	51.3	773	2	H20108
25	39	51.3	867	2	G69485
26	39	51.3	5037	1	A54151
27	38.5	50.7	391	2	AC0326
28	38	50.0	41	2	S77768
29	38	50.0	152	2	S68411

30	38	50.0	160	2	S68228	myeloid antimicrob
31	38	50.0	160	2	S68412	cathelin-related p
32	38	50.0	221	2	H90895	hypothetical prote
33	38	50.0	221	2	G85721	hypothetical prote
34	38	50.0	223	2	B96927	response regulator
35	38	50.0	270	2	T38878	DNAJ like protein
36	38	50.0	272	2	AB3430	transposase BME114
37	38	50.0	293	2	G64063	probable Mn transp
38	38	50.0	407	2	A99223	sugar phosphate nu
39	38	50.0	464	2	H82928	ATP synthase beta
40	38	50.0	547	2	B86723	NADH oxidase noxC
41	38	50.0	698	2	D64084	helicase (EC 3.6.1
42	38	50.0	854	2	AE1965	hypothetical prote
43	38	50.0	876	2	T52390	beta-1,4-xylosidas
44	38	50.0	985	1	VCLJSP	env polyprotein -
45	37.5	49.3	288	2	A81396	hypothetical prote

ALIGNMENTS

RESULT 1
S73462
transport ATP-binding protein pmcl - Mycoplasma pneumoniae (strain ATCC 29342)
N:Alternate names: hypothetical protein D12_01623
C:Species: Mycoplasma pneumoniae
A:Variety: ATCC 29342
C>Date: 26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 02-Feb-2001
C:Accession: S73462
R:Himmelfreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A:Reference number: S73327; MID:97105885
A:Accession: S73462
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-623 <HMM>
A:Cross-references: EMBL:AE000015; GB:U00089; NID:q1673779; PIDN:AA895784.1; PID:q167
C:Genetics:
A:Gene: pmcl
A:Genetic code: SGC3
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog
C:Keywords: ATP; nucleotide binding; P-loop
F:383-587/Domain: ATP-binding cassette homology <ABC>
F:400-407/Region: nucleotide-binding motif A (P-loop)

Query Match 57.9% Score 44; DB 2; Length 623;
Best Local Similarity 53.3% Pred. No. 21;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 RRIIRKIIHIKKNH 15
I:|||||:
DB 117 RKIRKIMHITAPSH 131

RESULT 2
A99114
hypothetical protein orf125 [imported] - Giardia theta nucleomorph
C:Species: nucleomorph Giardia theta
A>Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C:Accession: A99114
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.;
Nature 410, 1091-1096, 2001
A>Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671
A:Accession: A99114
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <DOU>
A:Cross-references: GB:AO10592; NID:q12580762; PIDN:CAC27080.1; GSPDB:GN00151

C:Genetics:
A:Gene: orf125
A:Map position: 2
A:Genome: nucleomorph
C:Keywords: nucleomorph

Query Match 55.3%; Score 42; DB 2; Length 125;
Best Local Similarity 70.0%; Pred. No. 9.7;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 RKIHIHIKK 13
|||||:
Db 9 RKIHIHLVK 18

RESULT 3
A33489
hypothetical protein, 36k - Anabaena sp. insertion sequence IS891
C:Species: Anabaena sp.

C>Date: 27-Feb-1990 #sequence_revision 27-Feb-1990 #text_change 02-Sep-2000
C:Accession: A33489

R:Bancroft, I.; Molk, C.P.
J. Bacteriol. 171, 5949-5954, 1989

A:Title: Characterization of an insertion sequence (IS891) of novel structure from the c
A:Reference number: A33489; MUID:90036679

A:Accession: A33489

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-310 <BAN>

A:Cross-references: GB:M24839; GB:M30792

A:Experimental source: strain M-131

C:Genetics:
A:Mobile element: insertion sequence IS891

C:Superfamily: hypothetical protein b1432

Query Match 55.3%; Score 42; DB 2; Length 310;
Best Local Similarity 70.0%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 RKIHIHIKK 14
|||||:
Db 95 RKVHALKKN 104

RESULT 4
T21413
hypothetical protein F26E4.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21413

R:Lightning, J.
submitted to the EMBL Data Library, October 1996

A:Reference number: Z19419

A:Accession: T21413

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-95 <WIL>

A:Cross-references: EMBL:Z81070; PIDN:CAB02999.1; GSPDB:GN00019; CESP:F26E4.6

A:Experimental source: clone F26E4

C:Genetics:

A:Gene: CESP:F26E4.6

A:Map position: 1

A:introns: 55/3

Query Match 53.9%; Score 41; DB 2; Length 95;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 RKIHIHIKK 15
|||||:
|::|||::|

Db 15 RQVIRPIVOAVRKG 29

RESULT 5
A10847

iron transport protein, periplasmic-binding protein sitA [Imported] - Salmonella ente

C:Species: Salmonella enterica subsp. enterica serovar Typh

A:Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: A10847

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr

S.; Moulé, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A:Reference number: AB0502; PMID:11677608

A:Accession: A10847

A:Molecule type: DNA

A>Status: preliminary

A:Residues: 1-305 <PAR>

A:Cross-references: GB:A1513382; PIDN:CAD05968.1; PID:q16503939; GSPDB:GN00176

C:Genetics:

A:Gene: sitA

C:Superfamily: adhesin B

Query Match 53.9%; Score 41; DB 2; Length 305;
Best Local Similarity 58.3%; Pred. No. 33;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKIHIHIKK 15
|||||:
Db 226 VRKVITFIKKHH 237

RESULT 6
S46010
hypothetical protein YBR141c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YBR1118

C:Species: Saccharomyces cerevisiae

C>Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 22-Oct-1999

C:Accession: S46010

R:Beckam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45995

A:Accession: S46010

A:Molecule type: DNA

A:Residues: 1-337 <BEC>

A:Cross-references: EMBL:Z36010; NID:q536439; PID:q536440; GSPDB:GN00002; MIPS:YBR141

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YBR141c

A:Map position: 2R

Query Match 53.9%; Score 41; DB 2; Length 337;
Best Local Similarity 61.5%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKIHIHIKK 13
|||||:
Db 30 RKIIRFHLINK 42

RESULT 7
F64056
probable ATP-dependent RNA helicase - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae

C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb-2001
C:Accession: F64056

R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
R.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman

, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: F64056
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-613 <TIGR>
 A:Cross-references: GB:U32709; GB:I42023; NID:g1573190; PIDN:AAC21900.1; PID:g1573195; T
 C:Keywords: ATP; nucleotide binding; P-loop
 F:49-56/Region: nucleotide-binding motif A (P-loop)
 F:151-156/Region: nucleotide-binding motif B
 F:155-158/Region: DEAD motif

Query Match 53.9%; Score 41; DB 2; Length 613;
 Best Local Similarity 53.8%; Pred. No. 65;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RRIIRKIIHIKK 13
 ||:| | | | |
 DB 353 RRLRNIEHLMKK 365

RESULT 8
 T18734
 Hypothetical protein B0391.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 R:Gardner, A.
 A:Submitted to the EMBL Data Library, November 1996
 A:Reference number: Z19012
 A:Accession: T18734
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1286 <ML>
 A:Cross-references: EMBL:Z81454; PIDN:GAB03805.1; GSPDB:GN00023; CESP:B0391.6
 A:Experimental source: clone B0391
 C:Genetics:
 A:Gene: CESP:B0391.6
 A:Map position: 5
 A:Introns: 64/2; 277/3; 350/2; 427/2; 494/2; 522/2; 749/3; 806/1; 872/2; 898/2; 1111/3;

Query Match 53.9%; Score 41; DB 2; Length 1286;
 Best Local Similarity 46.7%; Pred. No. 1.3e+02;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Oy 1 RRIIRKIIHIKKNH 15
 ||:| | | | |
 DB 526 KRVSRLKALVDKNH 540

RESULT 9
 G90478
 Hypothetical protein SS011614 [Imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: G90478
 R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, F.
 arrett, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.
 A:Submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: G90478
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-70 <RUR>
 A:Cross-references: GB:AE006641; NID:g13816367; PIDN:AAK43086.1; GSPDB:GN00155
 C:Genetics:

A:Gene: SS011614

Query Match 52.6%; Score 40; DB 2; Length 70;
 Best Local Similarity 63.6%; Pred. No. 12;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 4 IRKIIHIKKN 14
 :| | | | |
 DB 27 LRKIVEIKRN 37

RESULT 10
 C70464
 GTP-binding protein HflX - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Feb-2001
 C:Accession: C70464
 V:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 V.
 A:Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:9819666
 A:Accession: C70464
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-370 <AEF>
 A:Cross-references: GB:AE000762; NID:g2984163; PIDN:AAC07703.1; PID:g2984170; GB:AE00
 A:Experimental source: strain VF5
 C:Genetics:
 A:Gene: hflX
 C:Superfamily: GTP-binding protein hflX; translation elongation factor Tu homology
 C:Keywords: GTP binding; nucleotide binding; P-loop
 F:197-319/Domain: translation elongation factor Tu homology <ETD>
 F:203-210/Region: nucleotide-binding motif A (P-loop)
 F:316-319/Region: GTP-binding NKXD motif

Query Match 52.6%; Score 40; DB 2; Length 370;
 Best Local Similarity 61.5%; Pred. No. 58;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RRIIRKIIHIKK 13
 :| | | | |
 DB 159 KRLIKRIHIKK 171

RESULT 11
 T06357
 dnaK-type molecular chaperone Bip-C - soybean (fragment)
 N:Alternate names: binding protein Bip isoform C
 C:Species: Glycine max (soybean)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
 C:Accession: T06357
 R:Kalinski, A.; Rowley, D.L.; Loer, D.S.; Foley, C.; Buta, G.; Herman, E.M.
 A:Planta 195, 611-621, 1995
 A:Title: Binding-protein expression is subject to temporal, developmental and stress-
 A:Reference number: Z15624; MUID:95218610
 A:Accession: T06357
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-455 <KAL>
 A:Cross-references: EMBL:U08382; NID:g475597; PIDN:AA81953.1; PID:g475598
 A:Experimental source: Strain Century 84, Leaf
 C:Function:
 A:Description: Involved in protein folding and assembling/disassembling of protein co
 C:Superfamily: heat shock protein 70
 C:Keywords: ATP; molecular chaperone

Query Match 52.6%; Score 40; DB 2; Length 455;
 Best Local Similarity 53.3%; Pred. No. 71;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 RRIIRKIIHIKKNH 15
 :||: ||: ||| |
 Db 268 QRIMEYIIKLIKKNH 282

RESULT 12

S00180
 spermatid protein S1 - smaller spotted catshark
 C:Species: Scyliorhinus canicula (smaller spotted catshark, smaller spotted dogfish)
 C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 23-Feb-1997
 C:Accession: S00180
 R:Chaviviere, M.; Martindale, A.; Briand, G.; Sautiere, P.; Chevallerier, P.
 Eur. J. Biochem. 169, 105-111, 1987
 A:Title: Nuclear basic protein transition during sperm differentiation: amino acid sequ
 A:Reference number: S00180; MUID:88055030
 A:Accession: S00180
 A:Molecule type: protein
 A:Residues: 1-87 <CHA>
 C:Superfamily: sperm histone
 C:Keywords: DNA binding; nucleus; phosphoprotein; sperm
 F:3/55/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 51.3%; Score 39; DB 2; Length 87;
 Best Local Similarity 42.9%; Pred. No. 21;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRIIRKIIHIKKNH 14
 ||: |||: ||: ||:
 Db 42 RRRVRKIIHLKRRS 55

RESULT 13

H70318
 conserved hypothetical protein aq_202 - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999
 C:Accession: H70318
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666
 A:Accession: H70318
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-202 <AOF>
 A:Cross-references: GB:AE000678; NID:g2982921; PIDN:AA06551.1; PID:g2982934; GB:AE00065
 A:Experimental source: strain VFS
 C:Genetics: Aq_202
 A:Gene: aq_202
 C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0226

Query Match 51.3%; Score 39; DB 1; Length 202;
 Best Local Similarity 45.5%; Pred. No. 47;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 2 RIRKIIHIKK 12
 |: ||: ||: ||:
 Db 186 RALRLKLVHLK 196

RESULT 14

H89768
 hypothetical protein [imported] - Staphylococcus aureus (strain N315)
 C:Species: Staphylococcus aureus
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C:Accession: H89768
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Tuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hatori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
 A:Reference number: A89758; MUID:21311952; PMID:11418146
 A:Accession: H89768
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-203 <KUR>
 A:Cross-references: GB:BA000018; PID:g13700008; PIDN:BA041307.1; GSPDB:GN00149
 A:Experimental source: strain N315
 C:Genetics:
 A:Gene: SA0088

Query Match 51.3%; Score 39; DB 2; Length 203;
 Best Local Similarity 66.7%; Pred. No. 47;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 IIRKIIHIKKNH 14
 ||: ||: ||: ||:
 Db 25 IIEGIVHLKKNH 36

RESULT 15

F64620
 hypothetical protein HP0806 - Helicobacter pylori (strain 26695)
 C:Species: Helicobacter pylori
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
 C:Accession: F64620
 R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Feldman, J.M.; Fujii, C.; Bowman, C.; Wathey,
 Nature 388, 539-547, 1997
 A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser,
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467
 A:Accession: F64620
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-206 <TOM>
 A:Cross-references: GB:AE000592; GB:AE000511; NID:g2313929; PIDN:AA07859.1; PID:g231

Query Match 51.3%; Score 39; DB 2; Length 206;
 Best Local Similarity 53.8%; Pred. No. 48;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 IIRKIIHIKKNH 15
 ||: ||: ||: ||:
 Db 156 IIRKIIHIKKNH 168

Search completed: July 12, 2002, 07:56:53
 Job time: 220 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:04:38 ; Search time 46.15 Seconds
(without alignments)
12.585 Million cell updates/sec

Title: US-09-642-744B-15
Perfect score: 76
Sequence: 1 RRIIRKIHIIKKNH 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	57.9	623	Y014_MYCPN	P75095 mycoplasma
2	42	55.3	550	SGT1_SCHPO	Q9us49 schizosacch
3	41	53.9	337	YB21_YEAST	P38278 saccharomyc
4	41	53.9	475	VAT8_MANSE	Q9u5f0 manduca sex
5	41	53.9	613	DEAD_HAEIN	P44586 haemophilus
6	39	51.3	87	SSS1_SCYCA	P13275 scyllorhinu
7	39	51.3	88	Y113_METEX	O05113 methyllobact
8	39	51.3	431	RAM1_YEAST	P22007 saccharomyc
9	38	50.0	160	SC51_SHEEP	P49928 ovis aries
10	38	50.0	270	SC52_SHEEP	P49929 ovis aries
11	38	50.0	270	YD1J_SCHPO	O10247 schizosacch
12	38	50.0	293	Y362_HAEIN	Q57449 haemophilus
13	38	50.0	457	BAG4_HUMAN	O35429 homo sapien
14	38	50.0	670	REP_HAEIN	P44804 haemophilus
15	38	50.0	985	ENV_FOAMV	P14351 human spuma
16	37	48.7	211	ENGB_PVRAB	O5v288 pyrococcus
17	37	48.7	216	ENGB_PVRHO	O57938 pyrococcus
18	37	48.7	337	G3P1_AGABI	P32633 agaricus bi
19	37	48.7	352	YME7_YEAST	Q04705 saccharomyc
20	37	48.7	368	YSPB_BACSU	P44536 bacillus su
21	37	48.7	370	MATK_MARPO	P12174 marcanthia
22	37	48.7	571	PTL_CHLPN	Q62963 chlamydia p
23	37	48.7	682	GR78_YEAST	P16474 saccharomyc
24	37	48.7	788	PUR2_SCHPO	P20772 s difunctio
25	36	47.4	103	Y688_HAEIN	P44037 haemophilus
26	36	47.4	237	ISPD_BUCAI	P57475 buchnera ap
27	36	47.4	239	CENB_SHEEP	P9451 ovis aries
28	36	47.4	269	YHT4_YEAST	P38838 saccharomyc
29	36	47.4	277	SP0A_CLOIN	P32933 clostridium
30	36	47.4	289	Y552_HELPJ	Q9w158 helicobacte
31	36	47.4	305	MIAA_THEMEA	O9w127 drosophila
32	36	47.4	366	KORA_MERJA	O57724 methanococc
33	36	47.4	451	GD11_YEAST	P39956 saccharomyc

34	36	47.4	478	1	T23E_BACFE	Q02403 bacillus th
35	36	47.4	499	1	AMPA_BUCAI	P57448 buchnera ap
36	36	47.4	524	1	G6PD_DROME	P12646 drosophila
37	36	47.4	599	1	CENB_HUMAN	P07199 homo sapien
38	36	47.4	599	1	CENB_MOUSE	P27790 mus musculu
39	36	47.4	606	1	CENB_CRICR	P46988 cricetus
40	36	47.4	666	1	BIP_LYCES	P49118 lycopersico
41	36	47.4	667	1	BIP4_TOBAC	Q03684 nicotiana t
42	36	47.4	668	1	BIP5_TOBAC	Q03685 nicotiana t
43	36	47.4	687	1	NNP1_DROME	O9v127 drosophila
44	36	47.4	799	1	CN3B_MOUSE	O61409 mus musculu
45	36	47.4	817	1	YR86_CAEBL	Q09568 caenorhabdi

ALIGNMENTS

RESULT 1	ID	Y014_MYCPN	STANDARD:	PRT:	623 AA.
AC	P75095:				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Hypothetical ABC transporter ATP-binding protein MG014 homolog (D12_orf623).				
GN	MPN018 OR MP136.				
OS	Mycoplasma pneumoniae.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes; Mycoplasmataceae; Mycoplasma.				
OX	NCBI_TaxID=2104;				
XX	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 29342 / M129;				
RX	MEDLINE=97105885; PubMed=8948633;				
RA	Himmelfreith R., Hilbert H., Plagens H., Pirkl E., Li B.-C., Herrmann R.;				
RT	"Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";				
RL	Nucleic Acids Res. 24:4420-4449(1996).				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL; AE000015; AAB95784.1; -				
DR	HSSP; P1359; INED.				
DR	InterPro: IPR003593; AAA.				
DR	InterPro: IPR001140; ABC_transporter_tmam.				
DR	InterPro: IPR003439; ABC_transporter.				
DR	InterPro: IPR001687; ATP_GTP_A.				
DR	Pfam; PF00664; ABC_membrane; 1.				
DR	Pfam; SM00005; ABC_tran; 1.				
DR	SMART; SM00382; AAA; 1.				
DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.				
KW	Hypothetical protein; ATP-binding; Transport; Transmembrane; Complete proteome.				
KW	Complete proteome.				
FT	TRANSMEM 27				POTENTIAL.
FT	TRANSMEM 86				POTENTIAL.
FT	TRANSMEM 157				POTENTIAL.
FT	TRANSMEM 180				POTENTIAL.
FT	TRANSMEM 266				POTENTIAL.
FT	TRANSMEM 307				POTENTIAL.
FT	NP_BIND 400				ATP (POTENTIAL).
SO	SEQUENCE 623 AA; 68928 MW; AB0D481DBAEF736B CRC64;				

```

Query Match          57.9%; Score 44; DB 1; Length 623;
Best Local Similarity 53.3%; Pred. No. 4.8;
Matches      8; Conservative      3; Mismatches      4; Indels      0; Gaps      0;

OY      1 RRIIRKIIHIKKN 15
      1-1111111111
      117 RKIRKIMHITAPSH 131

Db

RESULT      2
SGTL_SCHPO
ID      SGTL_SCHPO      STANDARD;      PRT;      590 AA.
AC      Q9U549;
DT      01-MAR-2002 (Rel. 41, Created)
DT      01-MAR-2002 (Rel. 41, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      SgTl protein homolog C1002.10c.
GN      SPAC1002.10c.
OS      Schizosaccharomyces pombe (fission yeast).
OC      Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC      Schizosaccharomycetales; Schizosaccharomycetaceae;
OC      Schizosaccharomycetes.
OX      NCBI_TaxID=4896;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=972;
RA      Cadieu E., Lelaure V., Galibert F., McDougall R.C., Rajandream M.A.,
RA      Barrell B.G.,
RL      Submitted (JAN-2000) to the EMBL/GenBank/DBU databases.
-1- SIMILARITY: BELONGS TO THE SGTL FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; AJ136078; CAB65610.1; -.
KW      Hypothetical protein.
SQ      SEQUENCE      590 AA; 68156 MW; 1B0D7A283EB0294C CRC64;

Query Match          55.3%; Score 42; DB 1; Length 590;
Best Local Similarity 60.0%; Pred. No. 9.8;
Matches      6; Conservative      4; Mismatches      0; Indels      0; Gaps      0;

OY      5 RKIIHIKKN 14
      1111111111
      215 RKIVHVLQKN 224

Db

RESULT      3
YBZ1_YEAST
ID      YBZ1_YEAST      STANDARD;      PRT;      337 AA.
AC      P38278;
DT      01-OCT-1994 (Rel. 30, Created)
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      01-NOV-1995 (Rel. 32, Last annotation update)
DE      Hypothetical 38.5 kDa protein in IRA1-MAK5 intergenic region.
GN      YBR141C OR YBR1118.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX      NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=S288C;
RA      MEDLINE=95274325; PubMed=7754712;
RA      Zagulski M., Becam A.-M., Grzybowska E., Lacroute F., Migdalski A.,
RA      Sliominski P.P., Sokolowska B., Herbert C.J.;
RT      "The sequence of 12.5 kb from the right arm of chromosome II predicts

```

```

RT      a new N-terminal sequence for the IRA1 protein and reveals two new
RT      genes, one of which is a DEAD-box helicase."
RL      Yeast 10:1227-1234(1994).
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; Z36010; CAAB5099.1; -.
DR      EMBL; X78937; CAAS5538.1; -.
DR      PIR; S46010; S46010.
DR      SDD; S0000345; YBR141C.
KW      Hypothetical protein.
SQ      SEQUENCE      337 AA; 38539 MW; CB0039B18FAFB3BE CRC64;

Query Match          53.9%; Score 41; DB 1; Length 337;
Best Local Similarity 61.5%; Pred. No. 8.1;
Matches      8; Conservative      2; Mismatches      3; Indels      0; Gaps      0;

OY      1 RRIIRKIIHIKK 13
      1111111111
      30 RKIRRFHHLINK 42

Db

RESULT      4
VATH_MANSE
ID      VATH_MANSE      STANDARD;      PRT;      475 AA.
AC      Q9U5N0;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Vacuolar ATP synthase subunit H (EC 3.6.3.14) (V-ATPase H subunit).
DE      (Vacuolar proton pump H subunit) (Vacuolar proton pump subunit SFD).
OS      Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC      Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC      Sphingioidea; Sphingidae; Sphinginae; Manduca.
OX      NCBI_TaxID=7130;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Midgut;
RA      MEDLINE=20483512; PubMed=11030595;
RA      Merzendorfer H., Reineke S., Zhao X.F., Jacobmeier B., Harvey W.R.,
RA      Wiecekorek H.;
RT      "The multigene family of the tobacco hornworm V-ATPase: novel subunits
RT      a, C, D, H, and putative isoforms."
RL      Biochim. Biophys. Acta 1467:369-379(2000).
-1- FUNCTION: SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF VACUOLAR ATPASE.
-1- SUBUNIT H ACTIVATE ATPASE ACTIVITY OF THE ENZYME AND COUPLE ATPASE
ACTIVITY TO PROTON FLOW. VACUOLAR ATPASE IS RESPONSIBLE FOR
CELLS, THUS PROVIDING MOST OF THE ENERGY REQUIRED FOR TRANSPORT
PROCESSES IN THE VACUOLAR SYSTEM (BY SIMILARITY).
-1- SUBUNIT V-ATPase is an heteromultimeric enzyme composed of a
peripheral catalytic V1 complex (components A to H) attached to
an integral membrane V0 proton pore complex (components: a, c, c',
c'', and d).
-1- SIMILARITY: BELONGS TO THE V-ATPASE H SUBUNIT FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; AJ249389; CAB55499.1; -.

```

KM Hydrolase: ATP synthesis; Hydrogen ion transport.
SQ SEQUENCE 475 AA; 55024 MW; 39F3F7362305AE6 CRC64;

Query Match 53.9%; Score 41; DB 1; Length 475;
Best Local Similarity 26.7%; Pred. No. 12;
Matches 4; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRIIRKIIHIKKNH 15
Db 375 QELRTVHLLEKSH 389

RESULT 5
DEAD_HAEN ID DEAD_HAEN STANDARD: PRT: 613 AA.

AC P44586;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cold-shock DEAD-box protein A homolog (ATP-dependent RNA helicase dead homolog)
GN DEAD OR CSDA OR H10231.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX Haemophilus.
NCBI_TaxID=727;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischman R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA McKenney K., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Scott J.D., Shirley R., Liu L.-I., Glöckner A., Kelley J.M.,
RA Wetteman J.F., Phillips C.A., Spirys T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT Influenzae Rd.";
RL Science 269:196-512(1995).

CC -1- FUNCTION: HAS A HELIX-DESTABILIZING ACTIVITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: U32709; AAC21900.1; -;
DR HSP: Q58083; IHV8.
DR TIGR: H10231; -;
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000629; DEAD_ATP_helicase.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF002270; DEAD_1.
DR Pfam: PF002271; helicase_C_1.
DR SMART: SM00487; DEXDC_1.
DR SMART: SM00490; HELIC_C_1.
DR PROSITE: PS00039; DEAD_ATP_HELICASE_1.
KM Hydrolase; Helicase; ATP-binding; RNA-binding;
KM Transcription regulation; Complete proteome.
FT NP_BIND 49 56 ATP (POTENTIAL).
FT SITE 155 158 DEAD BOX.
SQ SEQUENCE 613 AA; 69705 MW; 1B826CBDEB1704DF CRC64;

Query Match 53.9%; Score 41; DB 1; Length 613;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRIIRKIIHIKK 13
Db 353 RRLRNIEHLMKK 365

RESULT 6
SSSL_SCYCA ID SSSL_SCYCA STANDARD: PRT: 87 AA.

AC P13275;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last annotation update)
DE Spermatid-specific protein S1.
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RX MEDLINE=88055030; PubMed=3678229;
RA Chauviere M., Martineau A., Briand G., Sautiere P., Chevallier P.,
RT "Nuclear basic protein transition during sperm differentiation. Amino
RT acid sequence of a spermatid-specific protein from the dog fish
RT Scyliorhinus caniculus.";
RL Eur. J. Biochem. 169:105-111(1987).
CC -1- FUNCTION: INVOLVED IN NUCLEAR BASIC PROTEIN TRANSITION: HISTONES
CC ARE REPLACED BY SPERMATID SPECIFIC PROTEINS WHICH ARE THEMSELVES
CC REPLACED BY PROTAMINES IN LATE SPERMATIDS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: WITH PROTEIN S2.
DR PIR: S00180; S00180.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Nuclear protein.
SQ SEQUENCE 87 AA; 11179 MW; 38A0ED7A82536910 CRC64;

Query Match 51.3%; Score 39; DB 1; Length 87;
Best Local Similarity 42.9%; Pred. No. 4;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 RRIIRKIIHIKKN 14
Db 42 RRRVRKIVHLKRS 55

RESULT 7
Y113_METEX ID Y113_METEX STANDARD: PRT: 88 AA.

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Hypothetical 9.7 kDa protein (ORF1).
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1 / NCIB 9133;
RX MEDLINE=97386438; PubMed=9244287;
RA Chistoserdova L., Lidstrom M.E.;
RT Identification and mutation of a gene required for glycinate kinase
RT activity from a facultative methylotroph, Methylobacterium extorquens
RT AM1.";
RL J. Bacteriol. 179:4946-4948(1997).
CC -1- SIMILARITY: STRONG, TO R.PROMAZEKII RP113.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U87316; AAB6493.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 88 AA: 9706 MW: C501A82AA35855F8 CRC64:

Query Match 51.3%; Score 39; DB 1; Length 88;
 Best Local Similarity 42.9%; Pred. No. 4.4;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 2 RIIRKIIHIKKNH 15
 :::|||:|::|
 Db 55 KVLKRIISLRKRDH 68

RESULT 8
 ID RAM1_YEAST STANDARD: PRT: 431 AA.
 AC P22007: O12422;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein farnesyltransferase beta subunit (EC 2.5.1.-) (CAAX
 DE farnesyltransferase beta subunit) (RAS proteins prenyltransferase
 DE beta) (Frase-beta).
 CN RAM1 OR DPR1 OR STR16 OR SCG2 OR YDL090C OR D2412.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89115553; PubMed=3064491;
 RA Goodman L.E., Perou C.M., Fujiyama A., Tamanoi F.;
 RT "Structure and expression of yeast PPR1, a gene essential for the
 RT processing and intracellular localization of ras proteins.";
 RL Yeast 4:271-281(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX Boskovic J., Salz J.E., Soler-Mira A., Garcia-Cantalejo J.M.,
 RA Ballesta J.P.G., Jimenez A., Remacha M.;
 RT "The sequence of a 16,691 bp segment of Saccharomyces cerevisiae
 RT chromosome IV identifies the DUN1, PMT1, PMT5, SRP14 and DPR1 genes,
 RT and five new open reading frames.";
 RL Yeast 12:1377-1384(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Wandut R., Medler H., Medler E., Scharfe M.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: CATALYZES THE TRANSFER OF A FARNESYL MOIETY FROM
 CC FARNESYL PYROPHOSPHATE TO PROTEINS SUCH AS A-FACTOR AND RAS
 CC PROTEINS.
 CC -I- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
 CC -I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
 CC -I- SIMILARITY: BELONGS TO THE PROTEIN PRENYLTRANSFERASE BETA SUBUNIT
 CC FAMILY.
 CC -I- SIMILARITY: CONTAINS 5 PFTB REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC or send an email to license@isb-sib.ch).

DR EMBL: M22753; AAA4579.1; ALT_SEQ.
 DR EMBL: X95644; CAA64921.1; -
 DR EMBL: Z74138; CAA98656.1; -
 DR PIR: S07864; BVBYDP.
 DR HSSP: 002293; 1FT1.
 DR SGD: S0002248; RAM1.
 DR InterPro: IPR001330; Prenyltrans.
 DR Pfam: PF00432; Prenyltrans; 5.
 KW Transferase; Prenyltransferase; Repeat; Zinc.
 FT REPEAT 130 171 PFTB 1.
 FT REPEAT 182 224 PFTB 2.
 FT REPEAT 231 273 PFTB 3.
 FT REPEAT 280 322 PFTB 4.
 FT REPEAT 332 375 PFTB 5.
 FT METAL 307 307 ZINC (BY SIMILARITY).
 FT METAL 309 309 ZINC (BY SIMILARITY).
 FT METAL 363 363 ZINC (BY SIMILARITY).
 FT CONFLICT 135 135 D -> V (IN REF. 1).
 SQ SEQUENCE 431 AA: 48189 MW: 2E3B64F30D2FF13A CRC64:

Query Match 51.3%; Score 39; DB 1; Length 431;
 Best Local Similarity 63.6%; Pred. No. 23;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 4 IRKIIHIKKN 14
 :|||:|
 Db 416 VKRIIHYFSN 426

RESULT 9
 ID SC51_SHEEP STANDARD: PRT: 160 AA.
 AC P49928;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Cathelin-related peptide SC5 precursor 1 (Antibacterial peptide SMAP-
 DE 29) (Myeloid antibacterial peptide SMAP-29).
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RX MEDLINE=96140581; PubMed=8549789;
 RA Mahoney M.M., Lee A.Y., Brezinski-Calliguri D.J., Huttner K.M.;
 RT "Molecular analysis of the sheep cathelin family reveals a novel
 RT antimicrobial peptide.";
 RL FEBS Lett. 377:519-522(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Huttner K.M., Mahoney M.M.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.
 CC -I- SIMILARITY: BELONGS TO THE CATHELICIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X92757; CAA63412.1; -
 DR EMBL: U06000; AAB49715.1; -
 DR InterPro: IPR001894; Cathelicidin.

```

DR Pfam: PF00666; Cathelicidins; 1.
DR ProDom: PD001838; Cathelicidin; 1.
DR PROSITE: PS00946; CATHELICIDIN_1; 1.
DR PROSITE: PS00947; CATHELICIDINS_2; 1.
KW Antibiotic; Signal.
FT SIGNAL 1 29
FT PROPEP 30 131
FT CHAIN 132 160
FT MOD_RES 30 30
FT DISULFID 86 97
FT DISULFID 108 125
SQ SEQUENCE 160 AA; 17786 MW; BD9B3859C432C249 CRC64;

Query Match
Best Local Similarity 61.5%; Score 38; DB 1; Length 160;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRIKKIHIHK 13
DB 135 RRLGRKIAHGK 147

RESULT 10
SC52_SHEEP STANDARD; PRT; 160 AA.
AC P49929;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cathelin-related peptide SC5 precursor 2 (Antibacterial peptide SMAP-29) (Weiloid antibacterial peptide SMAP-29).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=96140581; PubMed=8549789;
RA Mahoney M.M., Lee A.Y., Brezinski-Caliguri D.J., Hutter K.M.;
RT "Molecular analysis of the sheep cathelin family reveals a novel antimicrobial peptide.";
RL FEBS Lett. 377:519-522(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RX MEDLINE=96105386; PubMed=7498547;
RA Bagella L., Scocchi M., Zanetti M.;
RT "cDNA sequences of three sheep myeloid cathelicidins.";
RL FEBS Lett. 376:225-228(1995).
CC -1- FUNCTION: THERMOSTABLE, BROAD SPECTRUM, BACTERICIDAL AGENT.
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X92758; CAAG3413.1; -.
DR EMBL: I46854; AAA85470.1; -.
DR InterPro: IPR001894; Cathelicidin.
DR InterPro: IPR000010; Cystatin.
DR Pfam: PF00666; Cathelicidins; 1.
DR ProDom: PD001838; Cathelicidin; 1.
DR SMART: SM00043; CY; 1.
DR PROSITE: PS00946; CATHELICIDINS_1; 1.
DR PROSITE: PS00947; CATHELICIDINS_2; 1.

```

```

KW Antibiotic; Signal.
FT SIGNAL 1 29
FT PROPEP 30 131
FT CHAIN 132 160
FT MOD_RES 30 30
FT DISULFID 86 97
FT DISULFID 108 125
FT CONFLICT 28 28
SQ SEQUENCE 160 AA; 17742 MW; 4FB98A09355B51F CRC64;

Query Match
Best Local Similarity 61.5%; Score 38; DB 1; Length 160;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRIKKIHIHK 13
DB 135 RRLGRKIAHGK 147

RESULT 11
YDIJ_SCHPO STANDARD; PRT; 270 AA.
ID YDIJ_SCHPO
AC Q10247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 31.5 kDa protein C4G9.19 in chromosome I.
DE SPAC49.19.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A.,
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 J DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z69727; GAA93569.1; -.
DR HSSP: P25685; IHDJ.
DR InterPro: IPR001623; DnaJ_N.
DR Pfam: PF00226; DnaJ; 1.
DR SMART: SM00271; DnaJ; 1.
DR PROSITE: PS00636; DnaJ_1; 1.
DR PROSITE: PS0076; DnaJ_2; 1.
DR Hypothetical protein; Chapterone.
FT DOMAIN 43 112
SQ SEQUENCE 270 AA; 31520 MW; 37DE672DBA33C54A CRC64;

Query Match
Best Local Similarity 50.0%; Score 38; DB 1; Length 270;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 IIRKIIHIHKH 15
DB 48 IIRKIIELVKRH 60

RESULT 12
Y362_HAEIN

```



```

ID Y362_HAEIN STANDARD: PRT: 293 AA.
AC 057449; 005018;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Putative periplasmic iron-binding protein H10362 precursor.
GN H10362.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips R., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
RN [2]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RT "Two-dimensional map of the proteome of Haemophilus influenzae."
RL Electrophoresis 21:411-429(2000).
CC -I- FUNCTION: PART OF AN ATP-DRIVEN TRANSPORT SYSTEM
CC H10359/H10360/H10361/H10362 FOR IRON. METAL-BINDING COMPONENT
CC (POTENTIAL).
CC -I- SUBCELLULAR LOCATION: Periplasmic (potential).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL SOLUTE-BINDING PROTEIN FAMILY
CC 9.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32720; AAC22021.1; -
DR TIGR; H10362; -
DR InterPro: IPR001987; Lipoprotein_4.
DR Pfam: PF01297; Lipoprotein_4; 1.
DR PRINTS: PR00690; ADHESNFAMILY.
KW Transport; Periplasmic; Metal-binding; Iron; Signal;
KW Complete proteome.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 293 PUTATIVE PERIPLASMIC IRON-BINDING
FT PROTEIN H10362.
FT METAL 58 58 IRON (BY SIMILARITY).
FT METAL 123 123 IRON (BY SIMILARITY).
FT METAL 189 189 IRON (BY SIMILARITY).
FT METAL 189 189 IRON (BY SIMILARITY).
SO SEQUENCE 293 AA; 32463 MW; 27FE3F9BACDF508 CRC64;

Query Match 50.0%; Score 38; DB 1; Length 293;
Best Local Similarity 41.7%; Pred. No. 23;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
QY 4 IRKIIHIKKH 15
DB 220 VRKVIDLVKKN 231

```

```

RESULT 13
ID BAG4_HUMAN STANDARD: PRT: 457 AA.
AC 095429; 095818;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE BAG-family molecular chaperone regulator-4 (Silencer of death
DE domains).
DE BAG4 OR SODD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Carnivora; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99091615; PubMed=9873016;
RA Takayama S., Xie Z., Reed J.C.;
RT "An evolutionarily conserved family of Hsp70/Hsc70 molecular
RT chaperone regulators."
RL J. Biol. Chem. 274:781-786(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99115917; PubMed=9915703;
RA Jiang Y., Moronicz J.D., Liu W., Goeddel D.V.;
RT "Prevention of constitutive TNF receptor 1 signaling by silencer of
RT death domains."
RL Science 283:543-546(1999).
RN [3]
RP ERRATUM.
RA Jiang Y., Moronicz J.D., Liu W., Goeddel D.V.;
RL Science 283:1852-1852(1999).
CC -I- FUNCTION: INHIBITS THE CHAPERONE ACTIVITY OF HSP70/HSC70 BY
CC PROMOTING SUBSTRATE RELEASE.
CC -I- SUBUNIT: BINDS TO THE ATPASE DOMAIN OF HSP70/HSC CHAPERONES.
CC -I- SIMILARITY: CONTAINS 1 BAG DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF095194; AAD16123.2; -
DR EMBL: AF111116; AAD05226.1; -
DR MIM: 603884; -
DR InterPro: IPR003103; BAG.
DR Pfam: PF02179; BAG; 1.
DR SMART: SM00264; BAG; 1.
KW Chaperone.
FT DOMAIN 379 456 BAG.
SO SEQUENCE 457 AA; 49594 MW; B89D59E8118684A3 CRC64;

Query Match 50.0%; Score 38; DB 1; Length 457;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 4 IRKIIHIKK 13
DB 380 IRKIIHLEK 389

RESULT 14
ID REP_HAEIN STANDARD: PRT: 670 AA.
AC P44804;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```
DE ATP-dependent DNA helicase rep (EC 3.6.1.-).
GN REP OR HI0649.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fire L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd."
RL Science 269:496-512(1995).
CC -!- FUNCTION: REP HELICASE IS A SINGLE-STRANDED DNA-DEPENDENT
CC AT A NICK IN THE DNA. IT BINDS TO THE SINGLE-STRANDED DNA AND
CC ACTS IN A PROGRESSIVE FASHION ALONG THE DNA IN THE 3' TO 5'
CC DIRECTION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. UVRD SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32748; AAC22309.1; -.
DR HSSP: P09980; 10AA.
DR TIGR: HI0649; -.
DR InterPro: IPR000212; UVRD-helicase.
DR Pfam: PF00580; UVrd-helicase; 1.
KW Hydrolyase, Helicase; DNA replication; ATP-binding; DNA-binding;
KW Complete proteome.
FT NP_BIND 22 29 ATP (POTENTIAL);
FT SEQUENCE 670 AA; 77745 MW; 802A131251EDA24 CRC64;
```

```
Query Match 50.0%; Score 38; DB 1; Length 670;
Best Local Similarity 61.5%; Pred. No. 54;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 RRIIRKIHIKK 13
   |||||:|:|
DB 30 RVIINKIAHLIK 42
```

```
RESULT 15
ENV_FOAMV STANDARD; PRT; 985 AA.
ID ENV_FOAMV
AC P14351;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ENV polypeptide (Coat polypeptide).
GN ENV.
OS Human spumaretrovirus (Foamy virus).
OC Viruses; Retroid viruses; Retroviridae; Spumavirus.
OX NCBI_TaxId=11963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88004420; PubMed=2820721;
RA Fluegel R.M., Rethwilm A., Maurer B., Darai G.;
```

```
RT "Nucleotide sequence analysis of the env gene and its flanking
RT regions of the human spumaretrovirus reveals two novel genes."
RL EMBO J. 6:2077-2084(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X05591; CAA29086.1; -.
DR EMBL: M54978; AAA46123.1; -.
DR PIR: C29685; VCLJSP.
KW Coat protein; Transmembrane; Polypeptide; Glycoprotein.
FT TRANSMEM 64 84
FT TRANS 958 978
FT CARBOHYD 21 21
FT CARBOHYD 105 105
FT CARBOHYD 137 137
FT CARBOHYD 179 179
FT CARBOHYD 282 282
FT CARBOHYD 307 307
FT CARBOHYD 342 342
FT CARBOHYD 387 387
FT CARBOHYD 401 401
FT CARBOHYD 419 419
FT CARBOHYD 524 524
FT CARBOHYD 553 553
FT CARBOHYD 779 779
FT CARBOHYD 805 805
FT CARBOHYD 830 830
FT SEQUENCE 985 AA; 113494 MW; 60F588EA931BEOAF CRC64;
```

```
Query Match 50.0%; Score 38; DB 1; Length 985;
Best Local Similarity 33.3%; Pred. No. 80;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 1 RRIIRKIHIKKNH 15
   |||||:|:|
DB 677 KRIARSLVYVKQTH 691
```

```
Search completed: July 12, 2002, 08:04:39
Job time: 651 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:03:48 ; Search time 174.7 seconds
(without alignments)
14.854 Million cell updates/sec

Title: US-09-642-744b-15
Perfect score: 76
Sequence: 1 RRIIRKIHIIKKNH 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mmc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rident:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	55.3	125	10	Q9AVZ1
2	42	55.3	500	3	Q9P6R0
3	41	53.9	95	5	O18687
4	41	53.9	288	13	O91491
5	41	53.9	305	2	O9XCS2
6	41	53.9	311	2	O9X680
7	41	53.9	376	5	O9NDV0
8	41	53.9	390	5	O9NDU8
9	41	53.9	414	10	O947U2
10	41	53.9	787	5	O95XK1
11	41	53.9	1280	5	O9V255
12	40	52.6	70	17	O97UM2
13	40	52.6	370	16	O67743
14	40	52.6	455	10	O39803
15	40	52.6	477	5	O9VON8
16	40	52.6	523	5	O95TF9

17	39	51.3	202	16	O66580
18	39	51.3	203	16	O99XC0
19	39	51.3	206	16	O25486
20	39	51.3	206	16	O92L38
21	39	51.3	210	16	O92CL5
22	39	51.3	310	9	O9T171
23	39	51.3	310	16	O928X0
24	39	51.3	334	16	O9XOD8
25	39	51.3	407	3	O9GPM1
26	39	51.3	440	16	O97TNO
27	39	51.3	478	2	O9RGY1
28	39	51.3	483	16	O9KR81
29	39	51.3	497	2	O9FAB1
30	39	51.3	552	13	O9YGC1
31	39	51.3	665	3	O9HCO1
32	39	51.3	773	16	O51099
33	39	51.3	844	10	O94BN8
34	39	51.3	867	17	O28391
35	39	51.3	875	10	O9FLQ1
36	39	51.3	1136	10	O9LWG5
37	39	51.3	5037	13	O91313
38	38	50.0	441	2	O48971
39	38	50.0	147	11	O91VT5
40	38	50.0	164	17	O96YQ4
41	38	50.0	223	16	O97MH4
42	38	50.0	308	5	O9ND24
43	38	50.0	364	17	O9HIE5
44	38	50.0	407	17	O9UXD3
45	38	50.0	415	10	O9SPU4

ALIGNMENTS

RESULT 1	Q9AVZ1	PRELIMINARY:	PRT:	125 AA.
ID	Q9AVZ1;			
AC	Q9AVZ1;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	HYPOHEPTICAL 15.2 KDa PROTEIN.			
OS	Gulliardia theta (Cryptomonas ph1).			
OC	Eukaryota; Cryptophyta; Cryptomonadaceae; Gulliardia.			
OX	NCBI_TaxID=55529;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20087226; PubMed=10618395;			
RA	Zauner S., Fraunholz M., Wastl J., Penny S.L., Beaton M.,			
RA	Cavallier-Smith T., Maier U., Douglas S.;			
RT	"Chloroplast protein and centrosomal genes, a tRNA intron, and odd			
RT	telomeres in an unusually compact eukaryotic genome, the cryptomonad			
RT	nucleomorph.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:200-205(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21223349; PubMed=11323671;			
RA	Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.,			
RA	Wu X., Reith M., Cavallier-Smith T., Maier U.;			
RL	"The highly reduced genome of an enslaved algal nucleus.";			
DR	Nature 410:1091-1096(2001).			
EMBL	AF010592; CAC27080.1;			
KW	Hypothetical protein.			
SO	SEQUENCE 125 AA; 15207 MW; AAD538B5593997A1 CRC64;			
Query Match	55.3%;	Score 42;	DB 10;	Length 125;
Best Local Similarity	70.0%;	Pred. No. 18;		
Matches 7;	Conservative 3;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	4 IRKITHIKK 13			
Db	9 IRKITHIVRK 18			

```
RESULT 2
O9P6R0 PRELIMINARY; PRT; 500 AA.
ID O9P6R0;
AC O9P6R0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TRANSCRIPTION FACTOR IIB 70 KDA SUBUNIT.
GN SPBC13E7.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL545632, CAB89885.1; -.
DR InterPro: IPR004366; Cylin.
DR InterPro: IPR000812; TFIIB.
DR Pfam: PF00382; transcript_fac2. 2.
DR PRINTS: PR00685; TIFACTOR1IB.
DR SMART: SM00385; CYCLIN. 2.
SO SEQUENCE 500 AA; 56749 MW; 6CF1FD9ADA2B1B2 CRC64;

Query Match 55.3%; Score 42; DB 3; Length 500;
Best Local Similarity 41.7%; Pred. No. 66;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 1 RRIIRKIIHIK 12
II:::II:::
DB 247 RRSYREVYHYK 258

RESULT 3
O18687 PRELIMINARY; PRT; 95 AA.
ID O18687;
AC O18687;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE F26E4.6 PROTEIN.
GN F26E4.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lightning J.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology";
RL Science 282:2012-2018(1998).
DR EMBL: Z81070; CAB02999.1; -.
DR InterPro: IPR004202; COX7C.
DR Pfam: PF02935; COX7C. 1.
SO SEQUENCE 95 AA; 10828 MW; F0D82FB55C8D2314 CRC64;

Query Match 53.9%; Score 41; DB 5; Length 95;
Best Local Similarity 40.0%; Pred. No. 20;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
```

```
OY 1 RRIIRKIIHIK 15
II:::II:::
DB 15 RQIRPIVQAVRK 29

RESULT 4
O91491 PRELIMINARY; PRT; 288 AA.
ID O91491;
AC O91491;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE DYSTROPHIN (FRAGMENT).
OS Torpedo californica (Pacific electric ray).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hypnosquala; Pristigaster; Batoidae;
OC Torpediniformes; Torpedinidae; Torpedinidae; Torpedo.
OX NCBI_TaxID=7787;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ELECTRIC ORGAN;
RX MEDLINE=92291902; PubMed=1818109;
RA Ravin A.J., Dyer S.M., Yeaton J.E., Burden S.J.;
RT "Multiple dystrophin isoforms are associated with the postsynaptic
membrane of Torpedo electric organ.";
RL J. Physiol. (Paris) 85:131-133(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-ELECTRIC ORGAN;
RA Yeaton J.E.;
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: U01229; AAA03324.1; -.
DR InterPro: IPR002017; Spectrin.
FT NON_TER 1
SO SEQUENCE 288 AA; 32781 MW; E845F5CE5CAEC27C CRC64;

Query Match 53.9%; Score 41; DB 13; Length 288;
Best Local Similarity 42.9%; Pred. No. 57;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 2 RRIIRKIIHIK 15
II:::II:::
DB 183 KIINKLMDIKSH 196

RESULT 5
O9XCS2 PRELIMINARY; PRT; 305 AA.
ID O9XCS2;
AC O9XCS2;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SIRA.
GN SIRA.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SL1344;
RX MEDLINE=99185029; PubMed=10085045;
RA Zhou D., Hardt W.D., Galan J.E.;
RT "Salmonella typhimurium encodes a putative iron transport system
within the centisome 63 pathogenicity island.";
RL Infect. Immun. 67:1974-1981(1999).
DR EMBL: AF128999; AAD41065.1; -.
DR InterPro: IPR001987; Lipoprotein_4.
DR Pfam: PF01297; Lipoprotein_4; 1.
DR PRINTS: PR00690; ADHESINFAMILY.
SO SEQUENCE 305 AA; 33722 MW; 764EAE5332246FF0 CRC64;
```

Query Match 53.9%; Score 41; DB 2; Length 305;
 Best Local Similarity 58.3%; Pred. No. 61;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 IRKIITHIKKH 15
 :||:| |||:|
 Db 226 VRKVIDTIKHH 237

RESULT 6
 Q9X680 PRELIMINARY; PRT; 311 AA.
 AC Q9X680;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PUTATIVE PERIPLASMIC-BINDING PROTEIN.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SL1344;
 RA Damrauer S.D., Mills D.M., Lee C.A.;
 RT "A putative iron transporter is important for the growth and virulence
 of Salmonella typhimurium."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF132545; AAD30042.1; -;
 DR InterPro: IPR001987; Lipoprotein_4;
 DR Pfam: PF01297; Lipoprotein_4; 1.
 DR PRINTS; PR00690; ADHESNFAMILY.
 SQ SEQUENCE 311 AA; 34621 MW; 890FB36B887A4DA2 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 311;
 Best Local Similarity 58.3%; Pred. No. 62;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 IRKIITHIKKH 15
 :||:| |||:|
 Db 225 VRKVIDTIKHH 236

RESULT 7
 Q9NDV0 PRELIMINARY; PRT; 376 AA.
 AC Q9NDV0;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE B0391.6B PROTEIN.
 GN B0391.6B.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gardner A.E.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 DE none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; Z81454; CAB76406.1; -;
 DR InterPro: IPR002900; DUF38.
 DR Pfam; PF01827; F-box.
 DR SMART; SM00256; FBOX; 1.
 DR PROSITE; PS50181; FBOX; 1.
 SQ SEQUENCE 376 AA; 45527 MW; 0D84C52EB854299A CRC64;

DR Pfam; PF00646; F-box; 1.
 DR SMART; SM00256; FBOX; 1.
 DR PROSITE; PS50181; FBOX; 1.
 SQ SEQUENCE 376 AA; 44509 MW; 0AEFA0B1D90FCF7F CRC64;

Query Match 53.9%; Score 41; DB 5; Length 376;
 Best Local Similarity 46.7%; Pred. No. 74;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RRIIRKIITHIKKH 15
 ||:|:| :|||
 Db 86 RRVSRKLEALVCKNH 100

RESULT 8
 Q9NDU8 PRELIMINARY; PRT; 390 AA.
 AC Q9NDU8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE B0391.6A PROTEIN.
 GN B0391.6A.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gardner A.E.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL; Z81454; CAB76408.1; -;
 DR InterPro: IPR002900; DUF38.
 DR InterPro: IPR001810; F-box.
 DR Pfam; PF01827; DUF38; 1.
 DR Pfam; PF00646; F-box; 1.
 DR SMART; SM00256; FBOX; 1.
 DR PROSITE; PS50181; FBOX; 1.
 SQ SEQUENCE 390 AA; 45527 MW; 0D84C52EB854299A CRC64;

Query Match 53.9%; Score 41; DB 5; Length 390;
 Best Local Similarity 46.7%; Pred. No. 76;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 RRIIRKIITHIKKH 15
 ||:|:| :|||
 Db 100 RRVSRKLEALVCKNH 114

RESULT 9
 Q947U2 PRELIMINARY; PRT; 414 AA.
 AC Q947U2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE TRANSLATION INITIATION FACTOR 2B BETA SUBUNIT.
 GN OSJNBA0068A07.14.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=NIPPONBARE;
RA de la Bastide M., Spiegel L., Kirchoff K., Preston R., King L.,
RA Vil M.D., Baker J., Zutavern T., Santos L., Miller B., Kuit K.,
RA Cunniss D.M., Bell M., Balija V., Shah R., Bahret A., Dike S.,
RA Yang C., O'Shaughnessy A., Palmer L., Dedhia N., McCombie W.R.;
RT "Genomic Sequence for Oryza sativa, Nipponbare strain, clone
RT OSUNBA0068A07, from Chromosome 10, complete sequence.",
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC091734; AAK98757.1; -
KW Initiation factor.
SO SEQUENCE 414 AA; 44646 MW; BA54863017C7EF0C CRC64;

Query Match 53.9%; Score 41; DB 10; Length 414;
Best Local Similarity 54.5%; Pred. No. 81;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 3 RIIRKIIHIKK 13
|:|:|:|:|:|:|:
Db 76 IYRVLHIKE 86

RESULT 10
O95XFL PRELIMINARY; PRT; 787 AA.
ID O95XFL
AC O95XFL;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHEMETICAL 92.5 KDA PROTEIN.
GN Y73BA.15.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Leonard S.;
RT "The sequence of C. elegans cosmid Y73BA.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC084159; AAK39365.1; -
KW Hypothetical protein.
SO SEQUENCE 787 AA; 92494 MW; A27082E60A037935 CRC64;

Query Match 53.9%; Score 41; DB 5; Length 787;
Best Local Similarity 46.7%; Pred. No. 1.5e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 RIIRKIIHIKKNH 15
|:|:|:|:|:|:|:
Db 159 RVSRLKLVKDH 173

RESULT 11
O9VZ55 PRELIMINARY; PRT; 1280 AA.
ID O9VZ55

AC O9VZ55;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG1582 PROTEIN.
GN CG1582.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borzova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos R., Delcher A., Deltch A., Dey S., Dey J., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaislin D., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshire A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003484; AAF47973.1; -
DR FlyBase; FBgn030246; CG1582.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000571; zf-CCCH.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF00642; zf-CCCH; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR ATP-binding; Helicase.
SO SEQUENCE 1280 AA; 145120 MW; 6FC8661F4072B6EE CRC64;

Query Match 53.9%; Score 41; DB 5; Length 1280;
Best Local Similarity 53.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 RIIRKIIHIKKN 14
|:|:|:|:|:|:|:
Db 1267 RMIATVHLISKN 1279

RESULT 12
 0970M2
 ID 0970M2 PRELIMINARY; PRT; 70 AA.
 AC 0970M2;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE HYPOTHETICAL PROTEIN SS011614.
 GN SS011614.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=2132296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aweez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,
 RA Helkamp de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL: AE006889; AAA43086.1; -;
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 70 AA; 8403 MW; 7CDFFEBB8CAD9CD6 CRC64;

Query Match
 Best Local Similarity 52.6%; Score 40; DB 17; Length 70;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 4 IRRIIRIHIIRKN 14
 DB 27 LRKIYEIRKN 37

RESULT 13
 067743
 ID 067743 PRELIMINARY; PRT; 370 AA.
 AC 067743;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE GTP-BINDING PROTEIN HFLX.
 GN HFLX OR AO.1908.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Anjay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 397:353-358(1998).
 DR EMBL: AE000762; AAC07703.1; -;
 KM Complete proteome.
 SQ SEQUENCE 370 AA; 42100 MW; EBD137F19ABID215 CRC64;

Query Match
 Best Local Similarity 52.6%; Score 40; DB 16; Length 370;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RRIIRIHIIRKN 13

DB 159 KRLIKRIRHIIRK 171
 :|:|:| | | | | |
 RESULT 14
 039803
 ID 039803 PRELIMINARY; PRT; 455 AA.
 AC 039803;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE BIP ISOFORM C (FRAGMENT).
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CENTURY 84; TISSUE=LEAF;
 RX MEDLINE=95218610; PubMed=7766051;
 RA Kalinski A., Rowley D.L., Loefer D.S., Foley C., Buta G., Herman E.M.;
 RT "Binding-protein expression is subject to temporal, developmental and
 stress-induced regulation in terminally differentiated soybean
 organs.";
 RL Planta 195:611-621(1995).
 CC - SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
 DR EMBL: U08382; AAA81953.1; -;
 DR HSSP: P08107; 1H0.
 DR InterPro: IPR001023; HSP70.
 DR Pfam: PF00012; HSP70.1.
 DR PRINTS: PRO0301; HEATSHOCK70.
 DR PROSITE: PS00297; HSP70_1; 1.
 DR PROSITE: PS00329; HSP70_2; 1.
 DR PROSITE: PS01036; HSP70_3; 1.
 KM ATP-Binding.
 FT NON_TER
 SQ SEQUENCE 455 AA; 49824 MW; 10D120C35E3F1632 CRC64;

Query Match
 Best Local Similarity 52.6%; Score 40; DB 10; Length 455;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RRIIRIHIIRKNH 15
 DB 268 QRIMEYIRIKIRKKH 282

RESULT 15
 09VON8
 ID 09VON8 PRELIMINARY; PRT; 477 AA.
 AC 09VON8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CG3326 PROTEIN.
 GN CG3326.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Barton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT *The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003580; AAF51127.1; -;
 DR FlyBase: FBgn0031519; CG3326.
 DR InterPro: IPR003593; AAA.
 DR InterPro: IPR003959; AAA_subfam.
 DR Pfam: PF00004; AAA_1.
 DR SMART: SM00382; AAA; 1.
 KM ATP-binding.
 SQ SEQUENCE 477 AA; 53805 MW; 6EEB6784E133ED9E CRC64;

Query Match 52.6%; Score 40; DB 5; Length 477;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RRIIRKIHIIRKN 14
 : : | | : | | : | |
 Db 377 OKIIEKLIHQVKN 390

Search completed: July 12, 2002, 08:03:50
 Job time: 636 sec

xx Claim 10; Page 108; 137pp; English.
 PS
 CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
 CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
 CC SMP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
 CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
 CC antibiotic, antimicrobial and antiviral activities, and can be used as
 CC microbial growth and proliferation inhibitors and in gene therapy. (I)
 CC are useful for inhibiting microbial growth in an environment capable of
 CC sustaining such growth, for inhibiting microbial growth or strain in a
 CC host, and inhibiting the growth of drug-resistant microbial strains such
 CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
 CC *Xanthomonas*.
 CC
 SO Sequence 37 AA:
 Query Match 100.0%; Score 187; DB 22; Length 37;
 Best Local Similarity 100.0%; Pred. No. 3.5e-16;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GURRLRFRNKIKKIKGKIGGLPKLAPRTDY 37
 Db 1 glrklrkfrnkikkekikkgkigqllpklaprtdy 37
 |||||
 RESULT 2
 AAB51193 ID AAB51193 standard; Peptide; 38 AA.
 AC AAB51193;
 XX
 DT 22-MAR-2001 (first entry)
 XX
 DE E. coli AMP gene CAP18 amino acid sequence.
 XX
 KW Escherichia coli; E. coli; AMP gene; anti-microbial peptide;
 KW screening; preservation; food; feed; paint formulation; detergent;
 KW cosmetic; medical device; prosthetic implant; disinfectant;
 KW microbial infection; tumour.
 XX
 OS Escherichia coli.
 XX
 PN WO200073433-A1.
 PD 07-DEC-2000.
 XX
 PF 29-MAY-2000; 2000WO-DK00287.
 XX
 PR 31-MAY-1999; 99DK-0000766.
 XX
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Kristensen H;
 XX
 DR WPI; 2001-070965/08.
 XX
 PT Screening for nucleotide sequences encoding antimicrobial peptides by
 PT generating peptide libraries in microorganisms, inducing expression of
 PT peptides, selecting cells and recovering nucleotide sequences from
 PT cells -
 XX
 PS Example 1; Page 34; 59pp; English.
 CC The present invention describes a plasmid which is ligated with a pool
 CC of nucleotide sequences (NT) linked to an inducible promoter, to express
 CC a peptide (P) (an enzyme or mature (P) of less than 100 amino acids
 CC optionally linked to a signal (P)), transformed into host cells and
 CC cultured in presence of an inducer to induce expression of the NT. A
 CC method of screening (I) a pool of nucleotide sequences to select a
 CC nucleotide sequence encoding a peptide, comprises: (a) ligating a plasmid
 CC with the pool of NT; (b) transforming host cells which are sensitive to

CC the peptide with the ligated plasmids; (c) screening the transformed
 CC cells to select viable cells; (d) cultivating the viable cells in the
 CC presence of an inducer to induce expression of NT; (e) selecting cells
 CC according to the effect of the inducer on cell growth; and (f) recovering
 CC NT encoding the peptide from the selected cells. (I) is useful for
 CC screening a pool of nucleotide sequences to select a nucleotide sequence
 CC encoding a peptide which is an antimicrobial peptide or an antimicrobial
 CC enzyme active on bacteria and for finding and preparing a composition for
 CC treatment of human or animal. The antimicrobial peptide obtained using
 CC (I) may be employed in preservation of e.g. food/feed, paint
 CC formulations, detergents, cosmetics, medical devices such as prosthetic
 CC implants and also to disinfect and/or kill microbial cells on an object
 CC e.g. as an disinfectant for the treatment of biofilm. The peptides are
 CC useful for treating microbial infections and/or tumours. Peptides with
 CC improved bio-activity can be developed using (I). The peptides have no
 CC negative effect on normal mammalian and/or eukaryotic cells. The present
 CC sequence represents an AMP (antimicrobial peptide) gene amino acid
 CC sequence, which is used in an example from the present invention.
 CC
 SO Sequence 38 AA:
 Query Match 96.3%; Score 180; DB 22; Length 38;
 Best Local Similarity 94.6%; Pred. No. 2.6e-15;
 Matches 35; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GURRLRFRNKIKKIKGKIGGLPKLAPRTDY 37
 Db 2 glrklrkfrnkikkekikkgkigqllpklaprtdy 38
 |||||
 RESULT 3
 AAB28486 ID AAB28486 standard; peptide; 32 AA.
 AC AAB28486;
 XX
 DT 02-FEB-2001 (first entry)
 XX
 DE Rabbit lipopolysaccharide binding protein #1.
 XX
 KW Rabbit lipopolysaccharide binding protein #1.
 KW Lipopolysaccharide binding protein; lipopolysaccharide analyte;
 KW LPSA; immunoassay; gram-negative bacterium detection; Escherichia coli;
 KW Salmonella; Chlamydia; infection.
 XX
 OS Oryctolagus cuniculus.
 XX
 PN WO200060354-A1.
 PD 12-OCT-2000.
 XX
 PF 03-APR-2000; 2000WO-EP02869.
 XX
 PR 07-APR-1999; 99EP-0302711.
 XX
 PA (UNILE) UNILEVER PLC.
 PA (UNILE) UNILEVER NV.
 PA (HIND-) HINDUSTAN LEYER LTD.
 XX
 PI Badley RA, Hughes G, Zak KW;
 XX
 DR WPI; 2000-679390/66.
 XX
 PT Immunoassay for detecting lipopolysaccharides (LPS) from gram negative
 PT bacteria in a sample, comprises use of LPS-binding protein and an
 PT antibody to LPS as first or second binding reagents -
 XX
 PS Claim 12; Page 13; 40pp; English.
 CC The present sequence is lipopolysaccharide binding protein which may be
 CC used for detecting the presence of a lipopolysaccharide analyte (LPSA) in
 CC a sample. Anti-LPSA antibody is also used to detect the LPSA. The
 CC immunoassay is useful for detecting the presence of LPSA derived from the

CC cell membrane of a gram-negative bacterium such as *Escherichia coli*,
CC *Salmonella* and *Chlamydia*. It provides for assay devices to be
CC constructed which are capable of accurate and reproducible test
CC results, ideally suited for commercial markets such as the clinical or
CC home-testing markets. Furthermore, such immunoassays can be performed
CC quickly and simply, without using complex extracts or multiple assay
CC steps.

XX Sequence 32 AA;

Query Match 84.0%; Score 157; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.5e-12;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Oy 1 GLRRRLRFRNKKIKKIGKIGGLPKLA 32
Db 1 glrrrlrrfrnkikkkikgkigglpkla 32

RESULT 4

AAR45671
ID AAR45671 standard; Protein: 29 AA.

XX AAR45671;

XX 25-JUL-1994 (first entry)

XX N terminus of human CAP18.

XX Cationic antibacterial protein; lipopolysaccharide binding;
XX anticoagulant; granulocytes; RNIP; LPS; sepsis; autoimmune disorder;
XX septic shock; rabbit.

XX Homo sapiens.

XX Key Location/Qualifiers

XX FT Misc-difference 23 /label= Asp, Lys

XX FT Misc-difference 26 /label= Gln, Ile

XX FT Misc-difference 27 /label= Gly, Gln

XX MO9402589-A.

XX 03-FEB-1994.

XX 15-JUL-1993; 93WO-US06731.

XX 17-JUL-1992; 92US-0916761.

XX 17-JUL-1992; 92US-0916765.

XX (PANO-) PANORAMA RES INC.

XX Hirata M, Larrick JW, Wright SC;

XX WPI; 1994-048847/06.

XX Sequences encoding mammalian cationic antibacterial proteins -
XX are homologous to human and rabbit CAP18 sequences and have
XX lipopolysaccharide binding and anti-coagulation activity

XX Disclosure; Page 59; 112pp; English.

XX The sequence is that of a human cationic antibacterial protein CAP18
XX N-terminal fragment obtd. from granulocytes as sequenced. The
XX sequence corresponds to the N-terminus of the C-terminal RNIP fragment
XX of rabbit CAP18. The fragment is capable of binding to lipopolysaccharide
XX and inhibiting LPS-mediated activation of macrophage, as well as
XX interfering with the clotting cascade to inhibit coagulation in
XX conditions of disseminated intravascular coagulation. The polypeptides
XX can also be used to attenuate, inhibit or prevent LPS-associated

CC conditions, e.g., sepsis, autoimmune disorders, inflammation, etc.
CC See also AAR45667-81.

XX Sequence 29 AA;

Query Match 63.6%; Score 119; DB 15; Length 25;
Best Local Similarity 89.3%; Pred. No. 6.3e-08;
Matches 25; Conservative 0; Mismatches 3; Indels 0; Caps 0;

Oy 1 GLRRRLRFRNKKIKKIGKIGGL 28
Db 1 glrrrlrrfrnkikkkikgkiggl 28

RESULT 5

AAB07903
ID AAB07903 standard; Protein: 29 AA.

XX AAB07903;

XX 14-NOV-2000 (first entry)

XX Putative N-terminal of a human cationic protein designated CAP18.

XX Human; cationic protein; lipopolysaccharide binding; anticoagulant;
XX CAP18; reactive nitrogen inhibitory protein; RNIP; gram negative sepsis;
XX coagulation-related disorder; disseminated intravascular coagulation.

XX Homo sapiens.

XX Key Location/Qualifiers

XX FT Misc-difference 23 /label= Asp, Lys

XX FT Misc-difference 26 /label= Gln, Ile

XX FT Misc-difference 27 /label= Gly, Gln

XX US6103888-A.

XX 15-AUG-2000.

XX 01-JUN-1999; 99US-0322911.

XX 01-AUG-1996; 96US-0691280.

XX 17-JUL-1992; 92US-0916761.

XX 17-JUL-1992; 92US-0916765.

XX 27-SEP-1994; 94US-0313681.

XX (PANO-) PANORAMA RES INC.

XX Larrick JW, Wright SC, Hirata M;

XX WPI; 2000-531989/48.

XX Novel human cDNA encoding cationic proteins having lipopolysaccharide
XX binding and anticoagulant activity, useful for treating and diagnosing
XX gram negative sepsis and disseminated intravascular coagulation -

XX Disclosure; Column 11; 46pp; English.

XX The present sequence represents the putative N-terminal of a human
XX cationic protein, having lipopolysaccharide binding and anticoagulant
XX activity. The polypeptide is designated CAP18. Amino acids 134-170
XX of CAP18 represent a reactive nitrogen inhibitory protein (RNIP).
XX The CAP18 polynucleotide is useful for producing cationic proteins.
XX The CAP18 polypeptide is useful for treating and diagnosing
XX lipopolysaccharide-associated conditions such as gram negative
XX sepsis, and/or coagulation-related disorders, such as disseminated
XX intravascular coagulation.

SQ Sequence 29 AA;

Query Match

Best Local Similarity 63.6%; Score 119; DB 21; Length 29;
Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;OY 1 GLRKRLKFRNKIKKIKGKIGGL 28
DB 1 glrkrlkfrnkikikigkixxl 28
|||||

RESULT 6

AAB70668
ID AAB70668 standard; peptide; 22 AA.

AC AAB70668;

DT 15-MAY-2001 (first entry)

DE Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:21.

XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;

KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;

KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;

KM Burkholderia cepacia; Alcalligenes; Xanthomonas.

OS Oryctolagus cuniculus.

PN WO200112668-A1.

PD 22-FEB-2001.

PF 18-AUG-2000; 2000WO-US22781.

PR 18-AUG-1999; 99US-0149886.

PA (IOWA) UNIV IOWA RES FOUND.

PA (REGC) UNIV CALIFORNIA.

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

PS WPI; 2001-234911/24.

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
and proliferation of microbes, and for treating microbial infections

PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcalligenes and
CC Xanthomonas.

SQ Sequence 22 AA;

Query Match

Best Local Similarity 57.8%; Score 108; DB 22; Length 22;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;OY 3 RRLRKFRNKIKKIKGKIG 24
DB 1 rrlrkfrnkikikigkigki 22
|||||

RESULT 7

AAB70666
ID AAB70666 standard; peptide; 21 AA.

AC AAB70666;

DT 15-MAY-2001 (first entry)

DE Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:19.

XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;

KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;

KM proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;

KM Burkholderia cepacia; Alcalligenes; Xanthomonas.

OS Oryctolagus cuniculus.

PN WO200112668-A1.

PD 22-FEB-2001.

PF 18-AUG-2000; 2000WO-US22781.

PR 18-AUG-1999; 99US-0149886.

PA (IOWA) UNIV IOWA RES FOUND.

PA (REGC) UNIV CALIFORNIA.

PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;

PS WPI; 2001-234911/24.

PT New antimicrobial peptides useful as antibiotics for inhibiting growth
and proliferation of microbes, and for treating microbial infections

PS Claim 1; Page 103; 137pp; English.

CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcalligenes and
CC Xanthomonas.

SQ Sequence 21 AA;

Query Match 55.6%; Score 104; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 3 2e-06; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLRKRLKFRNKIKKIKG 21
DB 1 glrkrlkfrnkikikigki 21
|||||

RESULT 8

AAB70670
ID AAB70670 standard; peptide; 21 AA.

AC AAB70670;

DT 15-MAY-2001 (first entry)

DE Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:23.

XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;

KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;

```

XX (IOWA ) UNIV IOWA RES FOUND.
PA (REGC ) UNIV CALIFORNIA.
XX
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI: 2001-234911/24.
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
PS Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
XX Sequence 19 AA:
SQ
XX
XX Query Match 49.7%; Score 93; DB 22; Length 19;
XX Best Local Similarity 100.0%; Pred. NO. 6.5e-05;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 6 LRKFNKIKERLKKIGOKI 24
DB 1 LRKFNKIKERLKKIGOKI 19
XX
XX RESULT 10
XX AAB70667
XX ID AAB70667 standard; peptide: 18 AA.
XX
XX AAB70667;
XX
XX 15-MAY-2001 (first entry)
XX
XX Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:20.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Oryctolagus cuniculus.
XX
XX WO200112668-A1.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI: 2001-234911/24.
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
PS Claim 1; Page 103; 137pp; English.

```

XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 18 AA;

Query Match 47.6%; Score 89; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LRRKFRNKIKKIKKIG 21
Db 1 LRRKFRNKIKKIKKIG 18

RESULT 11
AAR94449
ID AAR94449 standard; peptide; 33 AA.
XX
AC AAR94449;
XX
DT 05-NOV-1996 (first entry)
XX
DE Mouse antimicrobial peptide ineffective syndecan expression inducer.
XX
KW Induction; expression; syndecan-1; syndecan-4; surface;
KW mesenchymal cell; fibroblast; epithelial; ineffective.
XX
OS Mus musculus.
XX
PN W09609322-A2.
XX
PD 28-MAR-1996.
XX
PF 22-SEP-1995; 95WO-US12080.
XX
PR 22-SEP-1994; 94US-0310722.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Bernfield M, Gallo RL;
XX
DR WPI; 1996-188401/19.
XX
PT Modulating mesenchymal interaction by administration of syndecan
PT used in the treatment of wounds, tumours, restenosis, etc
XX
PS Example 8; Page 27; 34pp; English.
XX
CC The present peptide, a mouse antibacterial peptide, is an
CC ineffective syndecan-1 and syndecan-4 mesenchymal cell, esp.
CC fibroblast and epithelial cell surface expression inducer.
CC Human microvascular endothelial cells were assayed for syndecan-4
CC expression following exposure to 5 % wound fluid, dbcAMP (1 mM),
CC the present peptide (10 microm) or a blank, to give respective
CC cell surface syndecan-4 values (MOD/m in) of approx. 1.75, 1.70,
CC 1.00 and 0.95.
XX
SQ Sequence 33 AA;

Query Match 43.9%; Score 82; DB 17; Length 33;
Best Local Similarity 58.1%; Pred. No. 0.0026;

Matches 18; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 6 LRRKFRNKIKKIKKIGQIGLLPKLAPRTD 36
Db 3 LRRKFRNKIKKIKKIGQIGLLPKLAPRTD 33

RESULT 12
AAR45669
ID AAR45669 standard; Protein; 37 AA.
XX
AC AAR45669;
XX
DT 25-JUL-1994 (first entry)
XX
DE N terminus of human CAP18.
XX
KW Cationic antibacterial protein; lipopolysaccharide binding;
KW anticoagulant; granulocytes; RNIP; LPS; sepsis; autoimmune disorder;
KW septic shock; rabbit.
XX
OS Homo sapiens.
XX
PN W09402589-A.
XX
PD 03-FEB-1994.
XX
PF 15-JUL-1993; 93WO-US06731.
XX
PR 17-JUL-1992; 92US-0916761.
XX
PR 17-JUL-1992; 92US-0916765.
XX
PA (PANO-) PANORAMA RES INC.
XX
PI Hirata M, Larrick JW, Wright SC;
XX
DR WPI; 1994-048647/06.
XX
PT Sequences encoding mammalian cationic antibacterial proteins -
PT are homologous to human and rabbit CAP18 sequences and have
PT Lipo:polysaccharide binding and anti-coagulation activity
XX
PS Claim 4; Page 58; 112pp; English.
XX
CC The sequence is that of a human cationic antibacterial protein CAP18
CC N-terminal fragment obtd. from granulocytes. The fragment is capable
CC of binding to lipopolysaccharide and inhibiting LPS-mediated activation
CC of macrophage, as well as interfering with the clotting cascade to
CC inhibit coagulation. The polypeptides can also be used to attenuate,
CC inhibit or prevent LPS-associated conditions, e.g. sepsis,
CC autoimmune disorders, inflammation, etc.
XX
SQ Sequence 37 AA;

Query Match 43.9%; Score 82; DB 15; Length 37;
Best Local Similarity 48.6%; Pred. No. 0.0029;
Matches 17; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 2 LRRKFRNKIKKIKKIGQIGLLPKLAPRTD 36
Db 2 LRRKFRNKIKKIKKIGQIGLLPKLAPRTD 36

RESULT 13
AAB07902
ID AAB07902 standard; Protein; 37 AA.
XX
AC AAB07902;
XX
DT 14-NOV-2000 (first entry)

```
PD      22-FEB-2001.
XX XX
PF      18-AUG-2000; 2000MO-US22781.
XX XX
PR      18-AUG-1999; 99US-0149886.
XX XX
PA      (IOWA ) UNIV IOWA RES FOUND.
RA      (REGC ) UNIV CALIFORNIA.
XX XX
PI      Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
DR XX
DP      WPL; 2001-234911/24.
XX XX
PT      New antimicrobial peptides useful as antibiotics for inhibiting growth
XX XX    and proliferation of microbes, and for treating microbial infections -
XX XX    Claim 1; Page 103; 137pp; English.
XX XX
CC      AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC CC    AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC CC    SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC CC    derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC CC    antibiotic, antimicrobial and antiviral activities, and can be used as
CC CC    microbial growth and proliferation inhibitors and in gene therapy. (I)
CC CC    are useful for inhibiting microbial growth in an environment capable of
CC CC    sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX XX
SQ      Sequence 17 AA:
XX XX
Query Match          43.3%; Score 81; DB 22; Length 17;
Best local similarity 100.0%; Pred. No. 0.0017;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      12 KIKELKKRIKGKIOGLL 28
        |||
        |KIKEKIKKIqkigll 17
Db
RESULT 15
AAR45673
ID      AAR45673 standard; Protein; 16 AA.
XX XX
AC      AAR45673;
XX XX
DT      25-JUL-1994 (first entry)
XX XX
DE      RNIP heparin binding fragment.
XX XX
KW      Cationic antibacterial protein; lipopolysaccharide binding;
XX XX    anticoagulant; granulocytes; RNIP; LPS; sepsis; autoimmune disorder;
KM      septic shock; rabbit; CAP18.
XX XX
OS      Synthetic.
XX OS
PN      WO9402589-A..
PP      03-FEB-1994.
XX PP
PF      15-JUL-1993; 93WO-US06731.
XX PF
PR      17-JUL-1992; 92US-0916761.
XX PR
PR      17-JUL-1992; 92US-0916765.
XX PR
PA      (PANO-) PANORAMA RES INC.
XX PA
EI      Hirata M, Larrick JW, Wright SC;
XX EI
WR      WPL; 1994-048847/06.
XX WR
```


PT Sequences encoding mammalian cationic antibacterial proteins -
PT are homologous to human and rabbit CAP18 sequences and have
PT lipo:polysaccharide binding and anti-coagulation activity
XX
PS Disclosure: Page 50; 112pp; English.
XX
CC The sequence of CAP18 C-terminal RNP was compared to that of a
CC number of heparin binding proteins to determine residues important for
CC binding to lipopolysaccharides and inhibiting LPS-mediated activation
CC of macrophage, as well as interfering with the clotting cascade to
CC inhibit coagulation in conditions of disseminated intravascular
CC coagulation.
CC See also AAR45667-81.
CC
XX
SQ Sequence 16 AA:

Query Match 42.8%; Score 80; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GIKRRLKKFRNKIKK 16
|||||
Db 1 gIKrIrIkfrnkIkK 16

Search completed: July 12, 2002, 08:00:43
Job time: 450 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:55:08 : Search time 75.52 Seconds
(without alignments)
11.967 Million cell updates/sec

Title: US-09-642-744B-17
Perfect score: 187
Sequence: 1 GLKRLRKFRNKKIKKIKGKIQGLLPKLAAPRTDY 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTOS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187	100.0	171	1	US-08-313-681A-4
2	187	100.0	171	3	US-09-322-911-4
3	119	63.6	29	1	US-08-313-681A-7
4	119	63.6	29	3	US-09-322-911-7
5	82	43.9	33	5	PCT-US95-12080-4
6	82	43.9	37	1	US-08-313-681A-6
7	82	43.9	37	3	US-09-322-911-6
8	80	42.8	16	1	US-08-313-681A-11
9	80	42.8	16	3	US-09-322-911-11
10	63	33.7	170	1	US-08-313-681A-2
11	63	33.7	170	3	US-09-322-911-2
12	57	30.5	30	1	US-07-725-331-62
13	57	30.5	30	5	PCT-US91-05047-62
14	57	30.5	36	1	US-07-725-331-63
15	57	30.5	36	5	PCT-US91-05047-63
16	52.5	28.1	77	1	US-08-182-175A-57
17	52.5	28.1	77	5	US-08-474-633A-75
18	52.5	28.1	77	5	PCT-US92-06412-57
19	51.5	27.5	381	4	US-08-983-035A-36
20	51.5	27.5	72	5	US-08-182-175A-87
21	51.5	27.5	72	5	PCT-US92-06412-87
22	51.5	27.5	107	1	US-08-182-175A-105
23	51.5	27.5	107	1	US-08-474-633A-92
24	51.5	27.5	107	5	PCT-US92-06412-105
25	51	27.3	236	3	US-08-493-071-15
26	51	27.3	277	1	US-08-690-457-5
27	51	27.3	277	2	US-08-628-187-5

28	51	27.3	277	3	US-08-493-071-3	Sequence 3, Appli
29	51	27.3	287	1	US-08-690-457-4	Sequence 4, Appli
30	51	27.3	287	2	US-08-628-187-4	Sequence 4, Appli
31	51	27.3	287	3	US-08-493-071-2	Sequence 2, Appli
32	51	27.3	288	1	US-08-690-457-3	Sequence 3, Appli
33	51	27.3	288	2	US-08-628-187-3	Sequence 3, Appli
34	51	27.3	288	3	US-08-493-071-1	Sequence 1, Appli
35	50	26.7	190	1	US-08-393-985-25	Sequence 25, Appli
36	50	26.7	263	1	US-08-393-985-23	Sequence 23, Appli
37	50	26.7	285	1	US-08-393-985-2	Sequence 2, Appli
38	50	26.7	288	1	US-08-337-602-4	Sequence 4, Appli
39	50	26.7	288	3	US-08-558-135-4	Sequence 4, Appli
40	50	26.7	288	4	US-08-819-286-3	Sequence 9, Appli
41	50	26.7	430	2	US-08-290-731C-9	Sequence 9, Appli
42	50	26.7	1572	2	US-08-290-731C-5	Sequence 5, Appli
43	49.5	26.5	1596	3	US-09-356-952-3	Sequence 3, Appli
44	49.5	26.5	28	1	US-08-182-175A-4	Sequence 4, Appli
45	49.5	26.5	28	1	US-08-182-175A-83	Sequence 83, Appli

ALIGNMENTS

```
RESULT 1
US-08-313-681A-4
; Sequence 4, Application US/08313681A
; Patent No. 5618675
; GENERAL INFORMATION:
; APPLICANT: Larlick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishimasa
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,681A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James W.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15325-9-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-313-681A-4

Query Match 100.0%, Score 187, DB 1, Length 171;
Best Local Similarity 100.0%, Pred. No. 8.2e-16;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 135 GLRRLRFRNKKIKKIKGKIGLPLKLAAPRTDY 171

RESULT 2

US-09-322-911-4
; Sequence 4, Application US/09322911
; Patent No. 6103888
; GENERAL INFORMATION:
; APPLICANT: Larrick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishimasa
; APPLICANT: Balint, Robert F.
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/322,911
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/691,280
; FILING DATE: August 1, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,761
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,765
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06731
; FILING DATE: July 15, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/313,681
; FILING DATE: September 27, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Flits, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 15325-000920
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-322-911-4

Query Match 100.0%; Score 187; DB 3; Length 171;

Best Local Similarity 100.0%; Pred. No. 8.2e-16;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GLRRLRFRNKKIKKIKGKIGLPLKLAAPRTDY 37

Db 135 GLRRLRFRNKKIKKIKGKIGLPLKLAAPRTDY 171

RESULT 3

US-08-313-681A-7
; Sequence 7, Application US/08313681A
; Patent No. 5618675
; GENERAL INFORMATION:
; APPLICANT: Larrick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishimasa
; APPLICANT: Balint, Robert F.
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,681A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15325-9-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 23
; OTHER INFORMATION: /note= "Xaa is Asp or Lys"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 26
; OTHER INFORMATION: /note= "Xaa is a Gln or Ile"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 27
; OTHER INFORMATION: /note= "Xaa is a Gly or Gln"
; US-08-313-681A-7

Query Match 63.6%; Score 119; DB 1; Length 29;

Best Local Similarity 89.3%; Pred. No. 2.4e-08;

Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GLRRLRFRNKKIKKIKGKIGLPL 28

Db 1 GLRRLRFRNKKIKKIKGKIGLPL 28

RESULT 4

US-09-322-911-7
; Sequence 7, Application US/09322911
; Patent No. 6103888
; GENERAL INFORMATION:
; APPLICANT: Larrick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishimasa
; APPLICANT: Balint, Robert F.

TITLE OF INVENTION: Human Cationic Proteins Having
TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/322,911
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/691,280
FILING DATE: August 1, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,761
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,765
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06731
FILING DATE: July 15, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/313,681
FILING DATE: September 27, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Filts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 15325-000920
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 23
OTHER INFORMATION: /note= "Xaa is Asp or Lys"
FEATURE:
NAME/KEY: Region
LOCATION: 26
OTHER INFORMATION: /note= "Xaa is a Gln or Ile"
FEATURE:
NAME/KEY: Region
LOCATION: 27
OTHER INFORMATION: /note= "Xaa is a Gly or Gln"
US-09-322-911-7

PCT-US95-12080-4
Sequence 4, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporation
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-12080-4

Query Match 43.9%; Score 82; DB 5; Length 33;
Best Local Similarity 58.1%; Pred. No. 0.00086;
Matches 18; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 6 LKRFNKIKKIKKIGOKIQLPRLAPRTD 36
DB 3 LKGGKIGKIKKIKKIKFQKLVQPE 33

RESULT 6
US-08-313-681A-6
Sequence 6, Application US/08313681A
Patent No. 5618675
GENERAL INFORMATION:
APPLICANT: Larlick, James W.
APPLICANT: Wright, Susan C.
APPLICANT: Hirata, Mishima
TITLE OF INVENTION: Human Cationic Proteins Having
TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,681A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541

REFERENCE/DOCKET NUMBER: 15325-9-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-313-681A-6

Query Match 43.9%; Score 82; DB 1; Length 37;
Best Local Similarity 48.6%; Pred. No. 0.00097;
Matches 17; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 2 LRRRLRRNRKIKKIKGKIGKIGLPLKPLARPTD 36
Db 2 LRRFRRSKEKIGKEKFRIVORIKDFLNLVPRTE 36

RESULT 7
US-09-322-911-6
Sequence 6, Application US/09322911
Patent No. 610388
GENERAL INFORMATION:
APPLICANT: Larrick, James W.
APPLICANT: Wright, Susan C.
APPLICANT: Hirata, Mishimasa
APPLICANT: Ballint, Robert F.
TITLE OF INVENTION: Human Cationic Proteins Having
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/322,911
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/691,280
FILING DATE: August 1, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,761
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,765
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06731
FILING DATE: July 15, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/313,681
FILING DATE: September 27, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 15325-000920
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-322-911-6

Query Match 43.9%; Score 82; DB 3; Length 37;
Best Local Similarity 48.6%; Pred. No. 0.00097;
Matches 17; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 2 LRRRLRRNRKIKKIKGKIGKIGLPLKPLARPTD 36
Db 2 LRRFRRSKEKIGKEKFRIVORIKDFLNLVPRTE 36

RESULT 8
US-08-313-681A-11
Sequence 11, Application US/08313681A
Patent No. 5618675
GENERAL INFORMATION:
APPLICANT: Larrick, James W.
APPLICANT: Wright, Susan C.
APPLICANT: Hirata, Mishimasa
APPLICANT: Ballint, Robert F.
TITLE OF INVENTION: Human Cationic Proteins Having
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,681A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15325-9-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-313-681A-11

Query Match 42.8%; Score 80; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLRRRLRRNRKIKK 16
Db 1 GLRRRLRRNRKIKK 16

RESULT 9

```
US-09-322-911-11
; Sequence 11, Application US/09322911
; Patent No. 6103888
; GENERAL INFORMATION:
; APPLICANT: Larlick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishimasa
; APPLICANT: Ballint, Robert F.
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/322,911
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/691,280
; FILING DATE: August 1, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,761
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,765
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06731
; FILING DATE: July 15, 1993
; APPLICATION NUMBER: 08/313,681
; FILING DATE: September 27, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 15325-000920
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-322-911-11

Query Match 42.8%; Score 80; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLRRRLKKFRKKIKK 16
Db 1 GLRRRLKKFRKKIKK 16

RESULT 10
US-08-313-681A-2
; Sequence 2, Application US/08313681A
; Patent No. 5618675
; GENERAL INFORMATION:
; APPLICANT: Larlick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishimasa
; APPLICANT: Ballint, Robert F.
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 08/313,681
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/691,280
; FILING DATE: August 1, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,761
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,765
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06731
; FILING DATE: July 15, 1993
; APPLICATION NUMBER: 08/313,681
; FILING DATE: September 27, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 15325-000920
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-322-911-11

Query Match 33.7%; Score 63; DB 1; Length 170;
Best Local Similarity 43.3%; Pred. No. 0.92;
Matches 13; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 7 RRRNRKIKKIKKIGKIGLLPKLAPRTD 36
Db 140 RRSKIKGKFRKIVIRIKDFLNLVPRTE 169

RESULT 11
US-09-322-911-2
; Sequence 2, Application US/09322911
; Patent No. 6103888
; GENERAL INFORMATION:
; APPLICANT: Larlick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishimasa
; APPLICANT: Ballint, Robert F.
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 08/313,681A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15325-9-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 170 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-313-681A-2
```

APPLICATION NUMBER: US/09/322,911
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/691,280
FILING DATE: August 1, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,761
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,765
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06731
FILING DATE: July 15, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/313,681
FILING DATE: September 27, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 15325-000920
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 170 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-322-911-2

Query Match 33.7%; Score 63; DB 3; Length 170;
Best Local Similarity 43.3%; Pred. No. 0.92;
Matches 13; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 7 RKRFRKIKKIKKIKKIGKIGLPLKAPRD 36
DB 140 RSKKIKKIKKIKKIKKIGLPLKAPRD 169

RESULT 12
US-07-725-331-62
Sequence 62, Application US/07725331
Patent No. 5294605
GENERAL INFORMATION:
APPLICANT: Houghten, Richard
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
TITLE OF INVENTION: Analogues Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
STREET: 180 No. 5294605th Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/725,331
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422

FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
OTHER INFORMATION: a C-terminal amide.
US-07-725-331-62

Query Match 30.5%; Score 57; DB 1; Length 30;
Best Local Similarity 43.3%; Pred. No. 0.85;
Matches 13; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 2 LKRLKFRKIKKIKKIKKIGKIGLPLK 31
DB 1 LKRLKRLKIKKIKKIKKIGLPLK 30

RESULT 13
PCT-US91-05047-62
Sequence 62, Application PC/TUS9105047

GENERAL INFORMATION:
APPLICANT: Houghten, Richard
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
TITLE OF INVENTION: Analogues Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
STREET: 180 North Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/05047
FILING DATE: 19910717
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide

FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
OTHER INFORMATION: a C-terminal amide.
PCT-US91-05047-62

Query Match 30.5%; Score 57; DB 5; Length 30;
Best Local Similarity 43.3%; Pred. No. 0.85;
Matches 13; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

OY 2 LKRLKFRNKIKKIKKIGKIGLLPKL 31
DB 1 LKRLKLLKLLKLLKLLKLLKLLKLL 30

RESULT 14
US-07-725-331-63
Sequence 63, Application US/07725331
Patent No. 5294605

GENERAL INFORMATION:
APPLICANT: Houghten, Richard
APPLICANT: Blondelle, Sylvie
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
TITLE OF INVENTION: Analogues Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
ADDRESSEE: 6 Milnamow
STREET: 180 No. 5294605th Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/725,331
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
OTHER INFORMATION: a C-terminal amide
US-07-725-331-63

Query Match 30.5%; Score 57; DB 1; Length 36;
Best Local Similarity 43.3%; Pred. No. 1;
Matches 13; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

OY 2 LKRLKFRNKIKKIKKIGKIGLLPKL 31
DB 7 LKRLKLLKLLKLLKLLKLLKLLKLL 36

RESULT 15
PCT-US91-05047-63
Sequence 63, Application PC/TUS9105047
GENERAL INFORMATION:

APPLICANT: Houghten, Richard
APPLICANT: Blondelle, Sylvie
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
TITLE OF INVENTION: Analogues Thereof
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Sutker, Shore,
ADDRESSEE: 6 Milnamow
STREET: 180 North Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/05047
FILING DATE: 19910717
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: acetylated at N-terminus, may be
OTHER INFORMATION: a C-terminal amide
PCT-US91-05047-63

Query Match 30.5%; Score 57; DB 5; Length 36;
Best Local Similarity 43.3%; Pred. No. 1;
Matches 13; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

OY 2 LKRLKFRNKIKKIKKIGKIGLLPKL 31
DB 7 LKRLKLLKLLKLLKLLKLLKLLKLL 36

Search completed: July 12, 2002, 07:55:08
Job time: 115 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:56:53 ; Search time 95.47 Seconds
(without alignments)
37.240 Million cell updates/sec

Title: US-09-642-744B-17
Perfect score: 187
Sequence: 1 GLRRLRKFRNKIKELKIKGKIQGLPLPLAPRTDY 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_71:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187	100.0	171	2 J01171	18K lipopolysaccha
2	82	43.9	173	2 S70521	cathelin-related p
3	65	34.8	324	2 T07982	probable choline-p
4	63	34.8	326	2 T07983	choline-phosphate
5	63	33.7	170	2 S74248	antibacterial pept
6	63	33.7	170	2 I38932	CAP18 precursor -
7	62.5	33.4	469	2 S69640	hypothetical prote
8	62	33.2	329	2 T07981	probable choline-p
9	62	33.2	331	2 T07980	probable choline-p
10	61.5	32.9	663	2 D97047	DNA ligase (NAD de
11	61.5	32.9	973	2 T50449	DNA repair and rec
12	60	32.1	121	2 H70471	conserved hypothet
13	59.5	31.8	307	2 B90236	transposase ISCI29
14	59	31.6	167	2 S68967	antibacterial pept
15	59	31.6	332	2 H84730	probable phospholi
16	58	31.0	3411	1 GNMVY	genome polypotein
17	58	31.0	3411	1 GNMVY	genome polypotein
18	57.5	30.7	401	2 A47141	apolipoprotein A-I
19	57.5	30.7	429	2 S29565	apolipoprotein A-I
20	57	30.5	342	2 E71101	probable geranylge
21	57	30.5	1163	1 GNMVY8	genome polypotein
22	57	30.5	3890	2 C89921	hypothetical prote
23	56.5	30.2	573	2 D90202	methionyl-trna syn
24	56	29.9	166	2 S41731	antibacterial prot
25	56	29.9	598	2 F12396	ABC transporter, A
26	55.5	29.9	1162	2 D83454	conserved hypothet
27	55.5	29.7	42	2 S65524	antibacterial prot
28	55	29.7	726	2 S73727	MG307 homolog H08
29	55	29.4	116	1 C64423	conserved hypothet

30	55	29.4	213	2 E83979	hypothetical prote
31	55	29.4	241	2 G86355	hypothetical prote
32	55	29.4	263	2 T30150	hypothetical prote
33	55	29.4	346	2 C75139	bifunctional short
34	55	29.4	354	2 B75003	hypothetical prote
35	55	29.4	498	2 T49333	related to mitotic
36	54.5	29.1	65	2 J00912	lys6 protein - yea
37	54.5	29.1	669	2 A97229	NAD-dependent DNA
38	54.5	29.1	1005	2 S19982	hypothetical prote
39	54	28.9	73	2 T03182	hypothetical prote
40	54	28.9	124	2 E82889	type I restriction
41	54	28.9	241	2 T27929	hypothetical prote
42	54	28.9	554	2 F86244	hypothetical prote
43	54	28.9	569	2 D69729	urease (alpha subu
44	54	28.9	1305	2 A40879	phospholipase C (E
45	54	28.9	1312	1 B40879	phospholipase C (E

ALIGNMENTS

```

RESULT 1
J01171
18K lipopolysaccharide-binding protein precursor - rabbit
M:Alternate names: 18K cationic protein
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 06-Dec-1996
C:Accession: J01171; PS0226
R:Larrick, J.W.; Morgan, J.G.; Pailings, I.; Hirata, M.; Yen, M.H.
Biochem. Biophys. Res. Commun. 179, 170-175, 1991
A:Title: Complementary DNA sequence of rabbit CAP18-a unique lipopolysaccharide bindi
A:Reference number: J01171; MUID:91354246
A:Accession: J01171
A:Molecule type: mRNA
A:Residues: 1-171 <LAR>
A:Experimental source: bone marrow
A:Accession: PS0226
A:Molecule type: protein
A:Residues: 135-159, 'GTGGLL' <LAZ>
A:Note: 157-Asp was also found
C:Superfamily: cathelin; cystatin homology
F:1-29/Domain: signal sequence #status predicted <Sig>
F:30-171/Product: 18K lipopolysaccharide-binding protein #status predicted <MAT>

Query Match 100.0%; Score 187; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 1.7e-14;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GLRRLRKFRNKIKELKIKGKIQGLPLPLAPRTDY 37
DB 135 GLRRLRKFRNKIKELKIKGKIQGLPLPLAPRTDY 171

RESULT 2
S70521
cathelin-related protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S70521
R:Popovska, A.E.; Zinovjeva, M.V.; Visser, J.W.M.; Zijlmans, J.M.J.M.; Fibbe, W.E.; B
FBBS Lett. 391, 5-8, 1996
A:Title: A novel murine cathelin-like protein expressed in bone marrow.
A:Reference number: S70521; MUID:96326596
A:Accession: S70521
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-173 <PO>
A:Cross-References: EMBL:X94353; NID:g1177533; PTDN:CA64078.1; PID:g1177534
C:Superfamily: cathelin; cystatin homology
F:1-23/Domain: signal sequence #status predicted <Sig>
F:24-144/Domain: propeptide #status predicted <PRO>
F:145-173/Product: cathelin-related protein #status predicted <MAT>

```

```

Query Match          43.9%; Score 82; DB 2; Length 173;
Best Local Similarity 58.1%; Pred. No. 0.023;
Matches 18; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY      6 LRKFRNKIKKKIGOKIOGLPKLPAPRTD 36
Db       142 LRRGGEKGELKKIGOKIKNFQKLVPQPE 172
      |||  || | ||||| || | :
      |||  || | ||||| || | :

RESULT      3
T07982
N:Alternate names: CTP:phosphocholine cytidylyltransferase
C:Species: Brassica napus (rape)
C>Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T07982
R:Nishida, I.; Swinhoe, R.; Slabas, A.R.; Murata, N.
Plant Mol. Biol. 31, 205-211, 1996
A>Title: Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase cDNAs by comp
A:Reference number: Z16266; MUID:96343925
A:Accession: T07982
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-324 <NIS>
A:Cross-references: EMBL:D03167; NID:d1049575; PIDN:BAA09643.1; PID:d1010288
A:Experimental source: cv. Jet Nurfi; roots
A:Genetics:
C:Function:
A:Description: catalyzes the synthesis of CDP-choline and PPI from CTP and phosphocholin
C:Keywords: nucleotidylyltransferase

Query Match          34.8%; Score 65; DB 2; Length 324;
Best Local Similarity 52.4%; Pred. No. 3.8;
Matches 11; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

OY      5 RLKRFNRKIKELKKIGOKIQ 25
Db       196 RLKLOEKVKKEQOEKVGEKIQ 216
      ||:| : |||: ||:||||
      ||:| : |||: ||:||||

RESULT      4
T07983
N:choline-phosphate cytidylyltransferase (EC 2.7.7.15) (clone CCT4) - rape
A:Alternate names: CTP:phosphocholine cytidylyltransferase
C:Species: Brassica napus (rape)
C>Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T07983
R:Nishida, I.; Swinhoe, R.; Slabas, A.R.; Murata, N.
Plant Mol. Biol. 31, 205-211, 1996
A>Title: Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase cDNAs by comp
A:Reference number: Z16266; MUID:96343925
A:Accession: T07983
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-326 <NIS>
A:Cross-references: EMBL:D63168; NID:d1049576; PIDN:BAA09644.1; PID:d1010289
A:Experimental source: cv. Jet Nurfi; roots
A:Genetics:
C:Function:
A:Description: catalyzes the synthesis of CDP-choline and PPI from CTP and phosphocholin
C:Keywords: nucleotidylyltransferase

Query Match          34.8%; Score 65; DB 2; Length 326;
Best Local Similarity 52.4%; Pred. No. 3.8;
Matches 11; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

OY      5 RLKRFNRKIKELKKIGOKIQ 25

```

```

||||:||||:||||:||||:
Db      198 RLKTKDEKVKKEQEKVGEKTK 218

RESULT      5
S74248
antibacterial peptide LL-37 precursor - human
N:Alternate names: antibacterial peptide FALL39
N:Contexts: antibacterial peptide LL-37; cathelin-related antibacterial peptide CAP-1
C:Species: Homo sapiens (man)
C:Date: 29-Jan-1998 #sequence-revision 06-Feb-1998 #text-change 16-Jul-1999
C:Accession: S74248; S78211; S66205; A55596; S49441; S58023
R:Gutmundsson, G.H.; Agerberth, B.; Odeberg, J.; Bergman, T.; Olsson, B.; Salcedo, R.
Eur. J. Biochem. 228, 325-332, 1996
A:Title: The human gene FALL39 and processing of the cathelin precursor to the antiba
A:Reference number: S74248; MUID:96283824
A:Accession: S74248
A:Molecule type: DNA
A:Residues: 1-170 <GUD>
A:Cross-references: EMBL:X96735
A:Accession: S78211
A:Molecule type: protein
A:Residues: 134-143 <GUE>
R:Cowland, J.B.; Johnsen, A.H.; Borregaard, N.
FEBS Lett. 368, 173-176, 1995
A:Title: hCAP-18, a cathelin/pro-bactenecin-like protein of human neutrophil specific
A:Reference number: S66205; MUID:95359966
A:Accession: S66281
A:Molecule type: mRNA
A:Residues: 1-112, 'T', 114-170 <COM>
A:Cross-references: EMBL:X89658; NID:g902627; PIDN:CAAC1805.1; PID:g902628
A:Accession: S66205
A:Molecule type: protein
A:Residues: 42-49;50-57;58-65, 'X', 67-85, 'X', 87-96, 'X', 98-100;110-112, 'T', 114-152 <COF
R:Agerberth, B.; Gunne, H.; Odeberg, J.; Kogner, P.; Bowman, H.G.; Gutmundsson, G.H.
Proc. Natl. Acad. Sci. U.S.A. 92, 195-199, 1995
A:Title: FALL-39, a putative human peptide antibiotic, is cysteine-free and expressed
A:Reference number: A55596; MUID:95116523
A:Accession: A55596
A:Molecule type: mRNA
A:Residues: 1-5, 'N', 7-112, 'T', 114-170 <AGE>
A:Cross-references: EMBL:Z38026; NID:g558378; PIDN:CAA6115.1; PID:g558379
C:Genetics:
A:Gene: FALL39
A:Map position: 3
A:Introns: 67/3; 103/3; 127/3
C:Superfamily: cathelin; cystatin homology
C:Keywords: antibacterial; blocked amino end; pyroglutamic acid
F:1-10/Domain: signal sequence #status predicted <SW>
F:21-130/Domain: cystatin homology <CR>
F:31-170/Product: cathelin-related antibacterial peptide CAP-18 #status predicted <MA
F:134-170/Product: propeptide #status predicted <PRO>
F:134-170/Product: antibacterial peptide LL-37 #status experimental <MAT>
F:31/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic

Query Match      33.7%; Score 63; DB 2; Length 170;
Best Local Similarity 43.3%; Pred. No. 3.5;
Matches 13; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY      7 RKFRNKIKELKKIKGKIOGLLPRLAPRTD 36
      11::11::11::11::11::11::11::
Db      140 RKSKEIKGEKFKRIVORIKDFLRLVPRTE 169

RESULT      6
I189932
CAP18 precursor - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence-revision 29-May-1998 #text-change 16-Jul-1999
C:Accession: I189932
R:Riattick, J.W.; Hirata, M.; Balint, R.F.; Lee, J.; Zhong, J.; Wright, S.C.
Infect. Immun. 63, 1291-1297, 1995

```

A:Title: Human CAP18: a novel antimicrobial lipopolysaccharide-binding protein.
A:Reference number: I38932; MUID:95197251
A:Accession: I38932
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-170 <RES>
A:Cross-references: EMBL:U0970; NID:9643476; PIDN:AAA74084.1; PID:9643477
C:Superfamily: cathelin; cystatin homology
F:22-130/Domain: cystatin homology <CYS>

Query Match 33.7%; Score 63; DB 2; Length 170;
Best Local Similarity 43.3%; Pred. No. 3.5;
Matches 13; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Oy 7 RKFRNKIKKIKGKIQGLLPKLPRTD 36
Db 140 RKSKEKIGKEFKRIVKQIKFLKNLVPRT 169

RESULT 7
S69640
hypothetical protein YDR473c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 23-Mar-2001
C:Accession: S69640
R:Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.
A:Reference number: S69554
A:Accession: S69640
A:Molecule type: DNA
A:Residues: 1-469 <DIE>
A:Cross-references: EMBL:U03050; NID:9927726; PID:9927760; GSPDB:GN00004; MIPS:YDR473c
C:Genetics:
A:Gene: MIPS:YDR473c
A:Map position: 4R

Query Match 33.4%; Score 62.5; DB 2; Length 469;
Best Local Similarity 46.9%; Pred. No. 10;
Matches 15; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

Oy 3 RKRFRNKIKKIKGKIQGLLPKLPAPR 34
Db 241 RKRFRNKIKKIKGKIQGLLPKLPAPR 271

RESULT 8
T07981
Probable choline-phosphate cytidylyltransferase (EC 2.7.7.15) (clone CCT1) - rape
N:Alternate names: CTP:phosphocholine cytidylyltransferase
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T07981
R:Nishida, I.; Swinhoe, R.; Slabas, A.R.; Murata, N.
Plant Mol. Biol. 31, 205-211, 1996
A:Title: Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase cDNAs by comp
A:Reference number: Z16266; MUID:96343925
A:Accession: T07981
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-329 <NIS>
A:Cross-references: EMBL:D58404; NID:d1044869; PIDN:BAA09571.1; PID:d1010214
A:Experimental source: cv. Jet Nurfi; roots
C:Genetics:
A:Gene: CCT
C:Function:
A:Description: catalyzes the synthesis of CDP-choline and PPI from CTP and phosphochol
C:Keywords: nucleotidyltransferase

Query Match 33.2%; Score 62; DB 2; Length 329;

Best Local Similarity 47.6%; Pred. No. 8.5;
Matches 10; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Oy 5 RLKFRNKIKKIKGKIQ 25
Db 205 RLKIQERVKKEQKVGKIQ 225

RESULT 9
T07980
Probable choline-phosphate cytidylyltransferase (EC 2.7.7.15) (clone CCT2) - rape
N:Alternate names: CTP:phosphocholine cytidylyltransferase
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T07980
R:Nishida, I.; Swinhoe, R.; Slabas, A.R.; Murata, N.
Plant Mol. Biol. 31, 205-211, 1996
A:Title: Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase cDNAs by c
A:Reference number: Z16266; MUID:96343925
A:Accession: T07980
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-331 <NIS>
A:Cross-references: EMBL:D63166; NID:d1049574; PIDN:BAA09642.1; PID:d1010287
A:Experimental source: cv. Jet Nurfi; roots
C:Genetics:
A:Note: CCT
C:Function:
A:Description: catalyzes the synthesis of CDP-choline and PPI from CTP and phosphoch
C:Keywords: nucleotidyltransferase

Query Match 33.2%; Score 62; DB 2; Length 331;
Best Local Similarity 47.6%; Pred. No. 8.5;
Matches 10; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Oy 5 RLKFRNKIKKIKGKIQ 25
Db 205 RLKIQERVKKEQKVGKIQ 225

RESULT 10
D97047
DNA ligase (NAD dependent), Liga [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: D97047
R:Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D97047
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-663 <KUP>
A:Cross-references: GB:AE001437; PIDN:AAK79167.1; PID:915024117; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1195
C:Superfamily: polynucleotidyltransferase (NAD+)

Query Match 32.9%; Score 61.5; DB 2; Length 663;
Best Local Similarity 46.4%; Pred. No. 19;
Matches 13; Conservative 6; Mismatches 6; Indels 3; Gaps 1;

Oy 10 RNKIKKIKGKIQGLLPKLPRTDY 37
Db 604 RTEIKKLESGARVAGSVSK---KTDY 628

RESULT 11

A:Map position: 2

Query Match 31.6%; Score 59; DB 2; Length 332;

Best Local Similarity 42.9%; Pred. No. 19;
Matches 9; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 5 RLKFRNKIKKKIGOKIO 25
||| : : : || : : : |||
Db 206 RLKFLQERVKKEOQERVKIO 226

Search completed: July 12, 2002, 07:56:55
Job time: 222 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:04:39 ; Search time 46.15 Seconds
(without alignments)
31.043 Million cell updates/sec

Title: US-09-642-744B-17
Perfect score: 187
Sequence: 1 GLRKRRLKFRNKIKELKIGKIGKIGLPLKLPRTDY 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	96.3	171	1	CP18_RABIT
2	82	43.9	173	1	GRAM_MOUSE
3	63	33.7	170	1	FA39_HUMAN
4	59	31.6	167	1	MP37_PIG
5	58	31.0	3411	1	POLG_YEVL
6	58	31.0	3411	1	POLG_YEVL
7	57.5	30.7	401	1	APR4_PAPAN
8	57.5	30.7	429	1	APR4_PAPAN
9	37	30.5	1163	1	POLG_YEVL
10	56	29.9	166	1	MP36_PIG
11	56	29.9	166	1	MP36_PIG
12	55.5	29.7	726	1	Y288_THEMA
13	55	29.4	113	1	YE40_MYCPN
14	54.5	29.1	65	1	Y516_YEVL
15	54.5	29.1	1005	1	Y516_YEVL
16	54	28.9	165	1	MB34_BOVIN
17	54	28.9	165	1	MB34_BOVIN
18	54	28.9	1111	1	URE1_BACSU
19	54	28.9	1312	1	Y1EP_HAEN
20	53.5	28.6	241	1	PIPI_DROME
21	53	28.3	188	1	YF58_ARATH
22	53	28.3	265	1	YF58_ARATH
23	53	28.3	288	1	APAI_PIG
24	53	28.3	366	1	ST1B_HUMAN
25	53	28.3	366	1	SPSY_MOUSE
26	52.5	28.1	1139	1	PARE_BORBU
27	52	27.8	425	1	KPCI_TRIPE
28	52	27.8	425	1	KPCI_TRIPE
29	51.5	27.5	229	1	Y537_BACND
30	51.5	27.5	266	1	REPX_STEAU
31	51.5	27.5	391	1	APR4_RAT
32	51.5	27.5	439	1	YB11_MERJA
33	51.5	27.5	477	1	SYC_PYRAB

34	51.5	27.5	551	1	MASY_CANTR	002216 candida tro
35	51.5	27.5	879	1	RA50_PYRHO	058687 pyrococcus
36	51.5	27.5	932	1	SYL_ARCFU	030250 archaeoglob
37	51	27.3	117	1	PFDB_PYRAB	09u74 pyrococcus
38	51	27.3	117	1	PFDB_PYRHO	058268 pyrococcus
39	51	27.3	259	1	1433_CHLRE	P52908 chlamydomon
40	51	27.3	265	1	APAI_BOVIN	P15497 bos taurus
41	51	27.3	564	1	MASY_BOVIN	P45458 glycine max
42	51	27.3	841	1	TRK1_SCHPO	P47946 schizosacch
43	51	27.3	1013	1	SBCC_VIBCH	09km67 vibrio chol
44	51	27.3	1027	1	KINN_MOUSE	P33175 mus musculu
45	51	27.3	1032	1	KINN_HUMAN	Q12840 homo sapien

ALIGNMENTS

RESULT ID	CP18_RABIT	STANDARD:	PRT:	171 AA.
AC	P25230:			
DT	01-MAY-1992 (rel. 22, Created)			
DT	01-MAY-1992 (rel. 22, Last sequence update)			
DT	15-JUL-1998 (rel. 36, Last annotation update)			
DE	Antimicrobial protein CAP18 precursor (18 kDa lipopolysaccharide-binding protein) (18 kDa cationic protein) (CAP18-A).			
CN	CAP18.			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 135-159.			
RC	TISSUE=Bone marrow;			
RX	MEDLINE=91354246; PubMed=1883348;			
RA	Larrick J.W., Morgan J.G., Pailings I., Hirata M., Yen M.H.;			
RT	"Complementary DNA sequence of rabbit CAP18 -- a unique lipopolysaccharide binding protein.";			
RL	Biochem. Biophys. Res. Commun. 179:170-175(1991).			
RN	[2]			
RP	SEQUENCE OF 135-159, AND CHARACTERIZATION.			
RX	MEDLINE=94178952; PubMed=8132348;			
RA	Hirata M., Shimomura Y., Yoshida M., Morgan J.G., Pailings I., Wilson D., Yen M.H., Wright S.C., Larrick J.W.;			
RT	"Characterization of a rabbit cationic protein (CAP18) with lipopolysaccharide-inhibitory activity.";			
RL	Intect. Immun. 62:1421-1426(1994).			
RN	[3]			
RP	SEQUENCE OF 135-154, AND CHARACTERIZATION.			
RX	MEDLINE=94075827; PubMed=8254193;			
RA	Larrick J.W., Hirata M., Zheng H., Zhong J., Bolin D., Cavallion J.W., Warren H.S., Wright S.C.;			
RT	"A novel granulocyte-derived peptide with lipopolysaccharide-neutralizing activity.";			
RL	J. Immunol. 152:231-240(1994).			
RN	[4]			
RP	CHARACTERIZATION.			
RX	MEDLINE=94148064; PubMed=8313956;			
RA	Tossi A., Scocchi M., Skerlavaj B., Gennaro R.;			
RT	"Identification and characterization of a primary antibacterial domain in CAP18, a lipopolysaccharide binding protein from rabbit leukocytes.";			
RL	FEBS Lett. 339:108-112(1994).			
RN	[5]			
RP	STRUCTURE BY NMR OF 135-166.			
RX	MEDLINE=95377455; PubMed=7649303;			
RA	Chen C., Brock R., Luh F., Chou P.-J., Larrick J.W., Huang R.-F., Huang T.-H.;			
RT	"The solution structure of the active domain of CAP18 -- a lipopolysaccharide binding protein from rabbit leukocytes.";			
RL	FEBS Lett. 370:46-52(1995).			
CC	-1- FUNCTION: CAP18 BINDS TO THE LIPID A MOIETY OF BACTERIAL LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER			


```

CC      MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. HAS ANTIBIOTIC ACTIVITY.
CC      -1- TISSUE SPECIFICITY: NEUTROPHILS.
CC      -1- SIMILARITY: BELONGS TO THE CATHHELICIDIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: M73998; AAA31187.1; -.
DR      PIR: J01171; J01171.
DR      PDB: 1LYP; 31-MAR-95.
DR      InterPro: IPR001894; Cathelicidin.
DR      Pfam: PF00666; Cathelicidins; 1.
DR      ProDom: PD001838; Cathelicidins; 1.
DR      PROSITE: PS00946; CATHHELICIDINS_1; 1.
DR      PROSITE: PS00947; CATHHELICIDINS_2; 1.
KW      Antibiotic; Signal; 3D-structure.
FT      SIGNAL 1 29 POTENTIAL.
FT      CHAIN 30 171 ANTIMICROBIAL PROTEIN CAP18.
FT      CHAIN 135 171 ANTIMICROBIAL PROTEIN CAP7.
FT      MOD_RES 30 30 PYROLIDONE CARBOXYLIC ACID
FT      (BY SIMILARITY).
FT      DISULFID 85 96 (BY SIMILARITY).
FT      DISULFID 107 124 BY SIMILARITY.
FT      VARIANT 157 157 K -> D.
SO      SEQUENCE 171 AA; 19805 MW; D7BF2103BCFB13C4 CRC64;

Query Match          96.3%; Score 180; DB 1; Length 171;
Best local Similarity 94.6%; Pred. No. 6.3e-14;
Matches 35; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 GLRRLRFRNRKIKKIKGKIGKIGLPLKLPRTDY 37
      |||||
Db      135 GLRRLRFRNRKIKKIKGKIGKIGFVPLKLPRTDY 171

RESULT 2
GRAM_MOUSE STANDARD: PRT: 173 AA.
AC      P51437;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      15-DEC-1998 (Rel. 37, Last annotation update)
DE      Cathelin-related antimicrobial peptide precursor (Cramp) (Cathelin-
DE      like protein) (CLP).
CN      CNLP OR CRAMP.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RX      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6; TISSUE=Bone marrow;
RX      MEDLINE=96326596; PubMed=8706528;
RA      Popsueva A.E., Zinovjeva M.V., Vissler Y.M.M., Fibbe W.E.,
RA      Belyavsky A.V.;
RT      "A novel murine cathelin-like protein expressed in bone marrow.";
RL      FEBS Lett. 391:5-8(1996).
RN      [3]

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=129/SVJ;
RA      Hutter K.M., Pirano J., Gallo R.L.;
RL      Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: ACTS AS A POTENT ANTIMICROBIAL PEPTIDE.
CC      -1- TISSUE SPECIFICITY: EXPRESSED IN TESTIS, SPLEEN, STOMACH, AND
CC      INTESTINE. VERY LOW EXPRESSION FOUND IN HEART, LUNG AND SKELETAL
CC      MUSCLE. NO EXPRESSION IN BRAIN, KIDNEY OR LIVER.
CC      -1- SIMILARITY: BELONGS TO THE CATHHELICIDIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: U43409; AA85898.1; -.
DR      EMBL: X94353; CA84078.1; -.
DR      EMBL: AF035680; AA88303.1; -.
DR      HSSP: P25230; 1LYP.
DR      MGI:108443; ClpP.
DR      InterPro: IPR001894; Cathelicidin.
DR      Pfam: PF00666; Cathelicidins; 1.
DR      ProDom: PD001838; Cathelicidins; 1.
DR      PROSITE: PS00946; CATHHELICIDINS_1; 1.
DR      PROSITE: PS00947; CATHHELICIDINS_2; 1.
KW      Antibiotic; Signal.
FT      SIGNAL 1 27 POTENTIAL.
FT      PROPEP 28 139 POTENTIAL.
FT      CHAIN 140 173 CATHELIN-RELATED ANTIMICROBIAL PEPTIDE.
FT      MOD_RES 28 28 PYROLIDONE CARBOXYLIC ACID (BY
FT      SIMILARITY).
FT      DISULFID 83 94 BY SIMILARITY.
FT      DISULFID 105 122 BY SIMILARITY.
FT      CONFLICT 24 24 L -> M (IN REF. 2).
FT      CONFLICT 173 173 MISSING (IN REF. 1).
SO      SEQUENCE 173 AA; 19581 MW; 143F0E84762F77E CRC64;

Query Match          43.9%; Score 82; DB 1; Length 173;
Best local Similarity 58.1%; Pred. No. 0.0099;
Matches 18; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy      6 LRRFRNRKIKKIKKIKGKIGKIGLPLKLPRTD 36
      |||  || |||||
Db      142 LRRGKGKIGKIKKIKGKIKNFQKIVPOPE 172

RESULT 3
FA39_HUMAN STANDARD: PRT: 170 AA.
AC      P49913;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Antibacterial protein FALL-39 precursor (FALL-39 peptide antibiotic)
DE      (Antimicrobial protein CAP-18) (LL-37).
GN      CAMP OR FALL39 OR CAP18.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RX      [1]
RP      SEQUENCE FROM N.A., AND SYNTHESIS OF 132-170.
RC      TISSUE=Bone marrow;
RX      MEDLINE=95116523; PubMed=7529412;
RA      Agerberth B., Gunne H., Odeberg J., Kogner P., Boman H.G.,
RA      Gudmundsson G.H.;
RT      "FALL-39, a putative human peptide antibiotic, is cysteine-free and
RT      expressed in bone marrow and testis.";
RL      Proc. Natl. Acad. Sci. U.S.A. 92:195-199(1995).
RN      [1]

```

```

Query Match      33.7%: Score 63; DB 1; Length 170;
Best Local Similarity 43.3%: Pred. No. 1.4;
Matches 13; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY      7  RKFRNKIKEXLKKIGOKTQGLLPKLAAPRTD 36
      ||::||::|:|:|:|:| | | |||:
Db      140  RKSEKIKGKFKRIVQRIKDFRLNLVPRT 169

RESULT      4
MP37_PIG

```

```

Query Match          31.6%; Score 59; DB 1; Length 167;
Best Local Similarity 45.2%; Pred. No. 4.1;
Matches 14; Conservative 7; Mismatches 6; Indels 4; Gaps 1;

QY      1 GLRRRLRFRN---KIKEIKLTKIGRIQGL 27
        || ||| | : : ||| ||||| : |
Db       131 GLSLRDPLSDRCGRRLGKEIKERIQKIKDL 161

RESULT    5
POLG_YEPV1 ID POLG_YEPV1 STANDARD: PRT: 3411 AA.
AC          P03J14; O42028;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [contains: Capsid protein C (Core protein); Matrix
DE protein (Envelope protein M); Major envelope protein E; Nonstructural
DE proteins NS1, NS2A, NS2B, NS4A and NS4B; Protease/helicase
DE (EC 3.4.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)
DE (NS5)]].
```

OS Yellow fever virus (strain 17D).
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 RX MEDLINE=85272570; PubMed=4023707;
 RA Rice C.M., Lencches E.M., Eddy S.R., Shin S.J., Sheets R.L.,
 RA Strauss J.H.;
 RT "Nucleotide sequence of yellow fever virus: implications for
 RT flavivirus gene expression and evolution.";
 RL Science 229:726-733(1985).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF FOUR PEPTIDE BONDS IN THE VIRAL
 CC PRECURSOR POLYPROTEIN, COMMONLY WITH ASP OR GLU IN THE P6
 CC POSITION, CYS OR THR IN P1 AND SER OR ALA IN P1'.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X03700; CAA27332.1; -
 DR PIR: A03914; GNMVY.
 DR HSP: A14336; ISVB.
 DR MEROPS: S07.001; -
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR000069; Flavi_M.
 DR InterPro: IPR001157; Flavi_NS1.
 DR InterPro: IPR000752; Flavi_NS2A.
 DR InterPro: IPR000487; Flavi_NS2B.
 DR InterPro: IPR000404; Flavi_NS4A.
 DR InterPro: IPR001528; Flavi_NS4B.
 DR InterPro: IPR000208; Flavi_NS5.
 DR InterPro: IPR001122; Flavi_capsid.
 DR InterPro: IPR000336; Flavi_glycoprote.
 DR InterPro: IPR001850; Flavi_glycoprote.
 DR InterPro: IPR002535; Flavi_helicase.
 DR InterPro: IPR002877; FtsJ.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF01003; Flavi_capsid.1.
 DR Pfam: PF00869; Flavi_glycoprote.1.
 DR Pfam: PF02832; Flavi_glycop_C.1.
 DR Pfam: PF00949; Flavi_helicase.1.
 DR Pfam: PF01004; Flavi_M.1.
 DR Pfam: PF00948; Flavi_NS1.1.
 DR Pfam: PF01005; Flavi_NS2A.1.
 DR Pfam: PF01002; Flavi_NS2B.1.
 DR Pfam: PF01350; Flavi_NS4A.1.
 DR Pfam: PF01349; Flavi_NS4B.1.
 DR Pfam: PF00972; Flavi_NS5.1.
 DR Pfam: PF01570; Flavi_propep.1.
 DR Pfam: PF01728; FtsJ.1.
 DR Pfam: PF00271; helicase_C.1.
 DR Pfam: PD001496; Flavi_NS1.1.
 DR Pfam: PD001556; Flavi_glycoprote.1.
 DR SMART: SM00450; HELIC_C.1.
 KW Polypeptide; Glycoprotein; Transferrase; RNA-directed RNA polymerase;
 KW Core protein; Coat protein; Envelope protein; Hydrolase; Helicase;
 KW ATP-binding; Transmembrane; Nonstructural protein.
 FT INT_MET 1
 FT CHAIN 1 121
 FT PROPEP 122 210 CAPSID PROTEIN C BY THE
 FT REMOVED AMINOPEPTIDASE.

FT CHAIN 211 285 ENVELOPE GLYCOPROTEIN M.
 FT CHAIN 285 778 MAJOR ENVELOPE PROTEIN E.
 FT CHAIN 778 1130 NONSTRUCTURAL PROTEIN NS1.
 FT CHAIN 1131 1354 NONSTRUCTURAL PROTEIN NS2A.
 FT CHAIN 1355 1484 NONSTRUCTURAL PROTEIN NS2B.
 FT CHAIN 1485 2107 PROTEASE/HELICASE (NS3).
 FT CHAIN 2108 2256 NONSTRUCTURAL PROTEIN NS4A.
 FT CHAIN 2257 2506 NONSTRUCTURAL PROTEIN NS4B.
 FT CHAIN 2507 3411 RNA-DIRECTED RNA POLYMERASE (NS5).
 FT CHAIN 3411 3411 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 271 285 POTENTIAL.
 FT TRANSMEM 271 285 POTENTIAL.
 FT TRANSMEM 740 753 POTENTIAL.
 FT TRANSMEM 755 778 POTENTIAL.
 FT TRANSMEM 1159 1180 POTENTIAL.
 FT DOMAIN 383 396 INVOLVED IN FUSION.
 FT NP_BIND 1682 1689 ATP (POTENTIAL).
 FT SITE 1773 1776 DEAD BOX.
 FT DISULFID 288 315 BY SIMILARITY.
 FT DISULFID 345 401 BY SIMILARITY.
 FT DISULFID 359 390 BY SIMILARITY.
 FT DISULFID 377 406 BY SIMILARITY.
 FT DISULFID 467 568 BY SIMILARITY.
 FT DISULFID 585 615 BY SIMILARITY.
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 986 986 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2320 2320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2346 2346 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2467 2467 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 3411 AA; 379512 MW; 680E0FACD23DCFA6 CRC64;

Query Match 31.0%; Score 58; DB 1; Length 3411.
 Best Local Similarity 36.4%; Pred. No. 75;
 Matches 12; Conservative 8; Mismatches 7; Indels 6; Gaps 1;
 Oy 2 LRRRLKFRNKIKKLTGOK-----:||: 28
 Db 16 VRRGVSLSKTKTKTKTGKRGVGGFT 48
 RESULT 6
 ID POLG_YEFV2 STANDARD: PRT; 3411 AA.
 AC P19901;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polypeptide [contains: Capsid protein C (Core protein); Matrix
 DE protein (Envelope protein M); Major envelope protein E; Nonstructural
 DE proteins NS1, NS2A, NS2B, NS4A and NS4B; Protease/helicase
 DE (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)
 DE (NS5)].
 OS Yellow fever virus (strain Pasteur 17D-204).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 ON NCBI_Taxid=11091;
 RX MEDLINE=89282413; PubMed=2734112;
 RA Dupuy A., Despres P., Cahour A., Girard M., Bouloy M.;
 RT "Nucleotide sequence comparison of the genome of two 17D-204 yellow
 RT fever vaccines.";
 RL Nucleic Acids Res. 17:3989-3989(1989).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF FOUR PEPTIDE BONDS IN THE VIRAL
 CC PRECURSOR POLYPROTEIN, COMMONLY WITH ASP OR GLU IN THE P6
 CC POSITION, CYS OR THR IN P1 AND SER OR ALA IN P1'.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL: X15062; CAB37419.1; -
DR PIR: S07757; GNMVYP.
DR HSSP: P14336; 1SVB.
DR MEROPS: S07.001; -
DR InterPro: IPR001410; DEAD
DR InterPro: IPR000069; Flavi_M.
DR InterPro: IPR001157; Flavi_NSI.
DR InterPro: IPR000752; Flavi_NS2A.
DR InterPro: IPR000487; Flavi_NS2B.
DR InterPro: IPR000404; Flavi_NS4A.
DR InterPro: IPR001528; Flavi_NS4B.
DR InterPro: IPR000208; Flavi_NS5.
DR InterPro: IPR001122; Flavi_capsid.
DR InterPro: IPR000336; Flavi_glycoprote.
DR InterPro: IPR001850; Flavi_helicase.
DR InterPro: IPR002535; Flavi_propep.
DR InterPro: IPR002877; FtsJ.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF01003; Flavi_capsid; 1.
DR Pfam: PF00869; Flavi_glycoprote; 1.
DR Pfam: PF02832; Flavi_glycop C; 1.
DR Pfam: PF00949; Flavi_helicase; 1.
DR Pfam: PF01004; Flavi_M; 1.
DR Pfam: PF00948; Flavi_NSI; 1.
DR Pfam: PF01005; Flavi_NS2A; 1.
DR Pfam: PF01002; Flavi_NS2B; 1.
DR Pfam: PF01350; Flavi_NS4A; 1.
DR Pfam: PF01349; Flavi_NS4B; 1.
DR Pfam: PF00972; Flavi_NS5; 1.
DR Pfam: PF01570; Flavi_propep; 1.
DR Pfam: PF01728; FtsJ; 1.
DR Pfam: PF00271; helicase_NSI; 1.
DR Pfam: PD001496; Flavi_NSI; 1.
DR ProDom: PD001556; Flavi_glycoprote; 1.
DR SMART: SM00490; HelicC; 1.
KW Polypeptide; Glycoprotein; Transferase; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Hydrolase; Helicase;
KW ATP-binding; Transmembrane; Nonstructural protein.
INT_MET 1 1

FT CHAIN 1 121
FT PROPEP 122 210
FT CHAIN 211 285
FT CHAIN 286 778
FT CHAIN 779 1130
FT CHAIN 1131 1354
FT CHAIN 1355 1484
FT CHAIN 1485 2107
FT CHAIN 2108 2256
FT CHAIN 2257 2506
FT CHAIN 2507 3411
FT NE_BIND 1682 1689
FT SITE 1773 1776
FT TRANSMEM 249 269
FT TRANSMEM 271 285
FT TRANSMEM 740 753
FT TRANSMEM 755 778
FT TRANSMEM 1159 1180
FT DISULFID 288 315
FT DISULFID 345 401
FT DISULFID 359 390
FT DISULFID 377 406

FT DISULFID 467 568 BY SIMILARITY.
FT DISULFID 585 645 BY SIMILARITY.
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 986 986 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2330 2330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2346 2346 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2467 2467 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 3411 AA; 379524 MW; 3298C0771FED23F7 CRC64;

Query Match 31.0%; Score 58; DB 1; Length 3411;
Best Local Similarity 36.4%; Pred. No. 75;
Matches 12; Conservative 8; Mismatches 7; Indels 6; Gaps 1;

QY 2 LRRRLKFRNKIKELKLTGK-----TGLL 28
Db 16 VRGVSLSNKIKOKTKOIGNRPGRVGQGT 48

RESULT 7
AP04_PAPAN STANDARD; PRT; 401 AA.
AC 028758;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Apolipoprotein A-IV precursor (Apo-AIV) (Fragment).
GN AP04.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OC NCBI_TaxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=93340170; PubMed=8101842;
RA Hixson J.E., Kammerer C.M., Molt G.E., Britten M.L., Birnbaum S.,
RA Powers P.K., Vandenberg J.L.;
RT "Baboon apolipoprotein A-IV. Identification of Lys-76-->Glu that
RT distinguishes two common isoforms and detection of length
RT polymorphisms at the carboxyl terminus.";
RL J. Biol. Chem. 268:15667-15673(1993).
CC -1- FUNCTION: MAY HAVE A ROLE IN CHYLOMICRONS AND VLDL SECRETION AND
CC CATABOLISM. REQUIRED FOR EFFICIENT ACTIVATION OF LIPOPROTEIN
CC LIASE BY APOC-II: POTENT ACTIVATOR OF LCAT. APOA-IV IS A MAJOR
CC COMPONENT OF HDL AND CHYLOMICRONS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED PRIMARILY IN THE INTESTINE.
CC SECRETED IN PLASMA.
CC -1- DOMAIN: NINE OF THE THIRTEEN 22-AMINO ACID TANDEM REPEATS (EACH
CC 22-MER IS ACTUALLY A TANDEM ARRAY OF TWO, A AND B, RELATED 11-
CC MERS) OCCURRING IN THIS SEQUENCE ARE PREDICTED TO BE HIGHLY ALPHA-
CC HELICAL, AND MANY OF THESE HELICES ARE AMPHIPHILIC. THEY MAY
CC THEREFORE SERVE AS LIPID-BINDING DOMAINS WITH LECTITHIN:CHOLESTEROL
CC ACYLTRANSFERASE (LCAT) ACTIVATING ABILITIES.
CC -1- POLYMORPHISM: THERE ARE TWO COMMON APOA-IV ISOFORMS, I (SHOWN
CC HERE) AND E. THE I ISOFORM IS ASSOCIATED WITH HIGHER LEVELS OF
CC HIGH DENSITY LIPOPROTEIN-C ON A HIGH CHOLESTEROL, SATURATED FAT
CC DIET.
CC -1- SIMILARITY: BELONGS TO THE APOA1 / APOA4 / APOE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L13174; AAA35379.1; -

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D1458; BAA03355.1; -
 DR PIR: J00374; GNMVY8.
 DR HSSP: P1436; ISVB.
 DR InterPro: IPR000069; Flavl_M.
 DR InterPro: IPR001157; Flavl_NSI.
 DR InterPro: IPR001122; Flavl_NSI.
 DR InterPro: IPR000336; Flavl_glycoprot.
 DR InterPro: IPR002535; Flavl_glycoprot.
 DR Pfam: PF01003; Flavl_capsid; 1.
 DR Pfam: PF00869; Flavl_glycoprot; 1.
 DR Pfam: PF02832; Flavl_glycoprot; 1.
 DR Pfam: PF01004; Flavl_M; 1.
 DR Pfam: PF00948; Flavl_NSI; 1.
 DR Pfam: PF01570; Flavl_propep; 1.
 DR Prodom: PD001496; Flavl_NSI; 1.
 DR Prodom: PD001556; Flavl_glycoprote; 1.
 DR Polyprotein: Glycoprotein: Coat protein;
 KM Envelope protein: Transmembrane; Nonstructural protein;
 FT INT_MET 1
 FT CHAIN 1 121
 FT PROPEP 122 210
 FT CHAIN 211 285
 FT CHAIN 286 778
 FT CHAIN 779 1130
 FT CHAIN 1131 >1163
 FT TRANSMEM 249 269
 FT TRANSMEM 271 285
 FT TRANSMEM 740 753
 FT TRANSMEM 755 778
 FT TRANSMEM 1159 >1163
 FT DISULFID 288 315
 FT DISULFID 345 401
 FT DISULFID 359 390
 FT DISULFID 377 406
 FT DISULFID 467 568
 FT DISULFID 585 615
 FT CARBOHYD 134 134
 FT CARBOHYD 150 150
 FT CARBOHYD 554 554
 FT CARBOHYD 908 908
 FT CARBOHYD 986 986
 FT NON_TER 1163 1163
 FT SEQUENCE 1163 AA; 128520 MW; 6BBDA317D722E01 CRC64;
 SO
 Query Match 30.5%; Score 57; DB 1; Length 1163;
 Best Local Similarity 36.4%; Pred. No. 38;
 Matches 12; Conservative 8; Mismatches 7; Indels 6; Gaps 1;
 Oy 2 LRKRLRFRNKIKKKIKGK-----IOGLL 28
 Db 16 VROGVRSLSNKKIKKIKGKRGVQGF 48
 RESULT 10
 MP36_PIG
 ID MP36_PIG STANDARD; PRT: 166 AA.
 AC P49931;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Antibacterial peptide PMAP-36 precursor (Myeloid antibacterial peptide
 DE 36)
 GN PMAP36.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OC NCB1_TaxID=9823;
 OX

RN [1]
 CC SEQUENCE FROM N.A., AND SYNTHESIS OF 130-166.
 RP TISSUE=Bone marrow;
 RC MEDLINE=94123775; PubMed=8293820;
 RX Stotici P., Scocchi M., Tossi A., Gennaro R., Zanetti M.
 RT "Chemical synthesis and biological activity of a novel antibacterial
 RT peptide deduced from a pig myeloid cDNA".
 RL FEBS Lett. 337:303-307(1994).
 CC -1- FUNCTION: EXERTS ANTIMICROBIAL ACTIVITY AGAINST BOTH GRAM-POSITIVE
 CC AND NEGATIVE BACTERIA. ITS ACTIVITY APPEARS TO BE MEDIATED BY ITS
 CC ABILITY TO DAMAGE BACTERIAL MEMBRANES.
 CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L29125; AAA31070.1; -
 DR InterPro: IPR001894; Cathelicidin.
 DR Pfam: PF00666; Cathelicidin; 1.
 DR Prodom: PD001838; Cathelicidin; 1.
 DR Prosite: PS00946; CATHELICIDINS_1; 1.
 DR Prosite: PS00947; CATHELICIDINS_2; 1.
 DR Antibiotic: Signal.
 FT SIGNAL 1 29
 FT PROPEP 30 129
 FT CHAIN 130 166
 FT MOD_RES 30 30
 FT DISULFID 85 96
 FT DISULFID 107 124
 FT SEQUENCE 166 AA; 18647 MW; 94B13C69709DA64B CRC64;
 SQ
 Query Match 29.9%; Score 56; DB 1; Length 166;
 Best Local Similarity 40.0%; Pred. No. 8.9;
 Matches 10; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
 Oy 5 RLKRLRFRNKIKKKIKGKIQGLP 29
 Db 132 RFRRLRKKIKKKIKGKYLKWI 156
 RESULT 11
 Y288_THEMA
 ID Y288_THEMA STANDARD; PRT: 598 AA.
 AC Q9WYQ4;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical ABC transporter ATP-binding protein TM0288.
 GN TM0288.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCB1_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99287316; PubMed=10360571;
 RX Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima".
 RL Nature 399:323-329(1999).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:03:50 ; Search time 174.7 Seconds
(without alignments)
36.639 Million cell updates/sec

Title: US-09-642-744B-17
Perfect score: 187
Sequence: 1 GURKRLRFRNKIKKKIGQKIGQLPKLAPRTDY 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP:REMBL_19:*
2: SP:archaea:*
3: SP:bacteria:*
4: SP:fungi:*
5: SP:human:*
6: SP_invertebrate:*
7: SP_mammal:*
8: SP_mhc:*
9: SP_organelle:*
10: SP_phage:*
11: SP_plant:*
12: SP_rodent:*
13: SP_virus:*
14: SP_vertebrate:*
15: SP_unclassified:*
16: SP_rvirus:*
17: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	43.3	170	6	O9GLV5
2	65	34.8	324	10	O42621
3	65	34.8	326	10	O42622
4	62.5	33.4	469	3	O03338
5	62	33.2	123	5	O9NHL6
6	62	33.2	329	10	O42620
7	62	33.2	331	10	O42619
8	61.5	32.9	663	16	O97J58
9	61.5	32.9	973	3	O9OR24
10	60	32.1	121	17	O67804
11	59.5	31.8	307	17	O97TW5
12	59	31.6	331	10	O42555
13	59	31.6	332	10	O92V56
14	59	31.6	406	16	O98J57
15	58	31.0	400	12	O89297
16	58	31.0	400	12	O89299

17	58	31.0	400	12	O89304	089304	yellow	feve
18	58	31.0	400	12	O89311	089311	yellow	feve
19	58	31.0	400	12	O89317	089317	yellow	feve
20	58	31.0	499	10	O91T20	O91T20	arabidops	is
21	58	31.0	778	12	O9W9B8	O9W9B8	yellow	feve
22	58	31.0	778	12	P89915	P89915	yellow	feve
23	58	31.0	3411	12	O98803	O98803	yellow	feve
24	58	31.0	3411	12	O91857	O91857	yellow	feve
25	58	31.0	3411	12	O9YMN2	O9YMN2	yellow	feve
26	58	31.0	3411	12	O9YWN1	O9YWN1	yellow	feve
27	58	31.0	3411	12	O9YWN0	O9YWN0	yellow	feve
28	58	31.0	3411	12	O9YRV3	O9YRV3	yellow	feve
29	58	31.0	3411	12	O89275	O89275	yellow	feve
30	58	31.0	3411	12	O89277	O89277	yellow	feve
31	58	31.0	3411	12	O89278	O89278	yellow	feve
32	58	31.0	3411	12	O89276	O89276	yellow	feve
33	57	30.5	400	12	O58799	O58799	pyrococcus	
34	57	30.5	400	12	O89295	O89295	yellow	feve
35	57	30.5	400	12	O89315	O89315	yellow	feve
36	57	30.5	3890	16	O99U53	O99U53	staphylococ	
37	56.5	30.2	413	10	O945H4	O945H4	hordeum	vul
38	56.5	30.2	573	17	O9UWM2	O9UWM2	sulfolobus	
39	56.5	30.2	598	2	O48537	O48537	lactobacill	
40	56	29.9	400	12	O89308	O89308	yellow	feve
41	56	29.9	400	12	O89309	O89309	yellow	feve
42	56	29.9	419	11	O9D5J9	O9D5J9	mus	musculu
43	56	29.9	1055	10	O9S722	O9S722	arabidops	is
44	56	29.9	1057	10	O9F117	O9F117	arabidops	is
45	56	29.9	1162	16	O91316	O91316	pseudomonas	

ALIGNMENTS

RESULT 1
O9GLV5
ID O9GLV5 PRELIMINARY: PRT: 170 AA.
AC O9GLV5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CATHELIN.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA Zhao C., Nguyen T., Lehrer R.I.;
RT "cDNA cloning of a monkey cathelin peptide.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF181954; AAC09440.1; -;
DR InterPro: IPR001894; Cathelin:clidin.
DR Pfam: PF00666; Cathelin:clidin; 1.
DR ProDom: PD001838; Cathelin:clidin; 1.
DR PROSITE; PS00946; CATHELICIDINS_1; 1.
SQ SQUIDNCE 170 AA: 18861 MW: 355AB3BF510DBB83 CRC64:

Query Match 43.3% Score 81; DB 6; Length 170;
Best Local Similarity 54.3% Pred. No. 0.041;
Matches 19; Conservative 3; Mismatches 9; Indels 4; Gaps 1;

Oy 5 RLRFRNRKIKK-----LKIGQKIGQLPKLAPRT 35
Db 134 RLGNFRKVRKIKGGGLKKYGVKTKDPLGNLVPT 168
RESULT 2
O42621
ID O42621 PRELIMINARY: PRT: 324 AA.

OY 2 LRKRLRK----FRNKIKKIKKIGOKIOG 26
 |||| : |||| : |||| : |||| : |||| :
 Db 13 LRKRRKKKTEFEREKLEPNLKKIKFOKIHG 41

RESULT 6
 Q42620 PRELIMINARY; PRT: 329 AA.

AC 042620;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CTP:PHOSPHOCHOLINE CYTIDYLYLTRANSFERASE (EC 2.7.7.15).
 GN CCT.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JET NURF; TISSUE=ROOT;
 RA Nishida I., Swincoe R., Slabas A.R., Murata N.;
 RT "Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase
 RT cDNAs by complementation of the in a yeast mutant."
 RL Plant Mol. Biol. 0:0-0(1996).
 DR EMBL: D58404; BAA09571.1; .
 DR HSSP: P27623; ICOZ.
 DR InterPro: IPR001994; Cytidylyltransf.
 DR Pfam: PF01467; Cytidylyltransf; 1.
 DR transferase; Nucleotidyltransferase.
 SQ SEQUENCE 329 AA; 38258 MW; 7A8A56D3CE703413 CRC64;

Query Match 33.2%; Score 62; DB 10; Length 329;
 Best Local Similarity 47.6%; Pred. No. 14;
 Matches 10; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

OY 5 LRKFRNKIKKIKKIGOKIO 25
 |||| : |||| : |||| : |||| : |||| :
 Db 205 LRKKLOERVKKEQOEKVGKIQ 225

RESULT 7
 Q42619 PRELIMINARY; PRT: 331 AA.

AC 042619;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CTP:PHOSPHOCHOLINE CYTIDYLYLTRANSFERASE (EC 2.7.7.15).
 GN CCT.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JET NURF; TISSUE=ROOT;
 RA Nishida I., Swincoe R., Slabas A.R., Murata N.;
 RT "Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase
 RT cDNAs by complementation of the in a yeast mutant."
 RL Plant Mol. Biol. 0:0-0(1996).
 DR EMBL: D63166; BAA09642.1; .
 DR HSSP: P27623; ICOZ.
 DR InterPro: IPR001994; Cytidylyltransf.
 DR Pfam: PF01467; Cytidylyltransf; 1.
 DR transferase; Nucleotidyltransferase.
 SQ SEQUENCE 331 AA; 38717 MW; CD487560667E2EEF CRC64;

Query Match 33.2%; Score 62; DB 10; Length 331;
 Best Local Similarity 47.6%; Pred. No. 14;
 Matches 10; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

OY 5 LRKFRNKIKKIKKIGOKIO 25
 |||| : |||| : |||| : |||| : |||| :
 Db 205 LRKKLOERVKKEQOEKVGKIQ 225

RESULT 8
 Q97JS8 PRELIMINARY; PRT: 663 AA.

AC 097JS8;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE DNA LIGASE (NAD DEPENDENT), LIGA.
 GN CAC1195.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Neolling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolff Y.I.,
 RA Tatusov R.L., Sabatne F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum."
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL: AE007633; AAK79167.1; .
 DR InterPro: IPR001357; BRCT.
 DR InterPro: IPR001679; DNA_ligase_N.
 DR InterPro: IPR004150; DNA_ligase_OB.
 DR InterPro: IPR004149; DNA_ligase_ZBD.
 DR InterPro: IPR003583; HHH_1.
 DR Pfam: PF00533; BRCT; 1.
 DR Pfam: PF01653; DNA_ligase_N; 1.
 DR Pfam: PF03120; DNA_ligase_OB; 1.
 DR Pfam: PF03119; DNA_ligase_ZBD; 1.
 DR ProDom: PD003944; DNA_ligase_N; 1.
 DR SMART: SM00292; BRCT; 1.
 DR SMART: SM00278; Hhh1; 3.
 DR SMART: SM00532; LIGANC; 1.
 DR PROSITE: PS0172; BRCT; 1.
 DR PROSITE: PS01056; DNA_LIGASE_N2; 1.
 DR Ligase; Complete proteome
 KW SEQUENCE 663 AA; 75209 MW; BB72425E32CC6959 CRC64;

Query Match 32.9%; Score 61.5; DB 16; Length 663;
 Best Local Similarity 46.4%; Pred. No. 31;
 Matches 13; Conservative 6; Mismatches 6; Indels 3; Gaps 1;

OY 10 RNKIKKIKKIGOKIGLLPKLAPRTDY 37
 |||| : |||| : |||| : |||| : |||| :
 Db 604 RTEIKKELKESGAKVAGSVSK---KTDY 628

RESULT 9
 Q9UR24 PRELIMINARY; PRT: 973 AA.

AC 09UR24;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE DNA REPAIR AND RECOMBINATION PROTEIN RHP26P.
 GN RHP26.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972H-;
 RC Brown S., Harris D., Lyne M.H., Rajandream M.A., Barrell B.G.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99410415; PubMed=10480889;
 RA Yasuhira S., Moriyomo M., Yasui A.;
 RT "transcription dependence and the roles of two excision repair
 RT pathways for UV damage in fission yeast Schizosaccharomyces pombe.";
 RL J. Biol. Chem. 274:26822-26827(1999).
 DR EMBL: AL133440; CAB62827.1; -
 DR EMBL: AB022912; BAAB4456.1; -
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR000330; SNF2_N.
 DR Pfam: PF00271; helicase_C; 1.
 DR Pfam: PF00176; SMF2_N; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELICC; 1.
 DR ATP-binding; Helicase.
 KW SEQUENCE 973 AA: 110913 MW: 4A8C61319B7D29EB CRC64;

Query Match:	32.9%	Score	61.5	DB	3	Length	973
Best Local Similarity	31.6%	Pred	No.	45			
Matches	12	Conservative	14	Mismatches	9	Indels	3
						Gaps	1

```

07      3 KRLRKRFRNKI---KEKLRKIGQKIGOLLPKLAPRTDY 378
      :||:| :| :|::: :| :|: :| :|
Db      41 KKRLLQKRKEISSYKEKIRLDERIDSRLLTKISYKKNF 788

```

RESULT	10
067804	
ID	067804
AC	067804
DT	01-AUG-1998 (TREMBLrel. 07, Created)
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE	HYPOTHETICAL 14.6 KDA PROTEIN.
GN	AQ_2000.
OS	Aquifex aeolicus.
OC	Bacteria; Aquificales; Aquificaceae; Aquifex.
OX	NCBI_TaxID=63363;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-VFS;
RA	MEDLINE=98196666; PubMed=9537320;
RA	Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA	Graham D.E., Overbeek R., Shead M.A., Keller M., Anjay M., Huber R.,
RA	Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT	"The complete genome of the hyperthermophilic bacterium Aquifex
RT	aeolicus";
RL	Nature 392:353-358(1998).
DR	EMBL: AE000767; AAC0774.1; -.
KM	Hypothetical protein; Complete proteome.
SO	SEQUENCE 121 AA; 14642 MW; C5ECFA762B0CD0C6 CRC64;

	Query Match	Similarity	32.1%	Score	60	DB	16	Length	121
	Best Local	Similarity	34.2%	Pred. No.	8.7				
	Matches	Conservative	13	Mismatches	9	Indels	12	Gaps	4
OY	4	KRLRFRRNKIKELKKI----	GKTIQGLLPKLAPRDI	37					
Db	28	KNLESFRKEIKSNKSNIQKVNSNEELOKIIEETKRRIIDY	65						

RESULT	11			
097TW5				
ID	097TW5	PRELIMINARY;	PRT;	307 AA.
AC	097TW5;			
DT	01-OCT-2001 (TREMblrel, 18, Created)			
DT	01-OCT-2001 (TREMblrel, 18, last sequence update)			
DT	01-OCT-2001 (TREMblrel, 18, last annotation update)			
DE	TRANSPOSASE ISCI290.			
GN	SSO1946 OR SSO0851 OR SSO1612.			
OS	Sulfolobus solfataricus.			
OC	Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.			
OX	NCBI_Taxid=2287;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 35092 / DSM 1617 / P2;			
RX	MEDLINE=21332296; Pubmed=11427726;			
RA	She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,			
RA	De Moers M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,			
RA	Ameyer A., Erruso G., Fletcher C., Gordon P.M.K.,			
RA	Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,			
RA	Thi-Nhoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,			
RA	Charlbois R.L., Doolittle W.F., Duguet M., Gaasterland T.,			
RA	Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.;			
RT	"The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).			
DR	EMBL; AE006802; AAK42139.1; -			
DR	EMBL; AE006707; AAK41145.1; -			
DR	EMBL; AE006774; AAK41827.1; -			
KW	Complete proteome.			
SQ	SEQUENCE 307 AA; 36660 MW; 27B9C5D321623C64 CHK64;			

Query Match	31.88;	Score	59.5;	DB	17;	Length	307;
Best Local Similarity	30.68;	Pred. No.	25;				
Matches	11;	Conservative	12;	Mismatches	10;	Indels	3;
						Gaps	1;

```
Qy      2 LKRLRKPRNKIKEKLKIGKIOGLPLKPLRTDY 37
      : : : : : | : : : : : : : : : |
Db      85 IHRRAKFRGEVKTLEKEYAKLEGGKMSRLA---DY 117
```

```

RESULT      12
042555
ID      Q42555      PRELIMINARY;      PRT;      331 AA.
AC      Q42555;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      CYP:PHOSPHOCHOLINE CYTIDYLYLTRANSFERASE (EC 2.7.7.15).
OS      Arabidopsis thaliana (mouse-ear cress).
OC      Eudicotyledons; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatopsida; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eustroids II; Brassicales; Brassicaceae; Arabidopsids.
RX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=COLUMBA1A;
RA      Choi S.-B., Lee K.-W., Cho S.-H.;
RL      submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR      EMBL; U50451; AAA93035.1; -.
DR      HSSP; P27623; IC02.
DR      InterPro: IPR001994; Cytidylyltransf.
DR      Pfam; PF01467; Cytidylyltransf. 1.
DR      Transferase; Nucleotidyltransferase.
SQ      SEQUENCE      331 AA;      38417 MW;      85D7B9594E4E0B04 CMC64;

```

```

Query Match          31.6%;  Score 59;  DB 10;  Length 311;
Best Local Similarity 42.9%;  Pred. NO. 31;
Matches 9;  Conservative 9;  Mismatches 3;  Indels 0;  Gaps 0;

0y      5 RLKRFNRNKKKKKKIGOKIO 25
      |||:| : :|||: :|||:|

```

Db 205 RLKLLQERKVEQDERVGEKIQ 225

RESULT 13

09ZV56 PRELIMINARY: PRT: 332 AA.

AC 09ZV56: PRELIMINARY: PRT: 332 AA.

DT 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE PUTATIVE PHOSPHOCHOLINE CYTIDYLTRANSFERASE.

GN T32F6.22 OR CCT.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RN SEQUENCE FROM N.A.

RP STRAIN=CV. COLUMBIA;

RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,

RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;

RT "Arabidopsis thaliana chromosome II BAC T32F6 genomic sequence.";

RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA COL-O;

RA Choi Y.H., Choi S.B., Cho S.H.;

RT "Structure of a CTP:phosphocholine cytidyltransferase Gene from

RL Arabidopsis thaliana";

RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.

DR EMBL: AC005700; AAC69950.1; -

DR EMBL: AF165912; AAD45922.1; -

DR HSSP: P27623; ICDZ.

DR InterPro: IPR001994; Cytidylyltransf.

DR Pfam: PF01467; Cytidylyltransf. 1

DR Transferase; Nucleotidyltransferase.

KW SEQUENCE 332 AA; 38485 MW; 536E2FB87907AB54 CRC64;

SO

Query Match 31.6%; Score 59; DB 10; Length 332;

Best local Similarity 42.9%; Pred. No. 31;

Matches 9; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 5 RLKFRNKIKKIKGKIQ 25

Db 206 RLKLLQERKVEQDERVGEKIQ 226

RESULT 14

098J57 PRELIMINARY: PRT: 406 AA.

AC 098J57: PRELIMINARY: PRT: 406 AA.

DT 01-OCT-2001 (TREMBlrel. 18, Created)

DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)

DE 01-OCT-2001 (TREMBlrel. 18, Last annotation update)

DE MR2092 PROTEIN.

GN MR2092.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RP STRAIN=MAFF303099;

RX MEDLINE=21082930; PubMed=11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti";

RL DNA Res. 7:331-338(2000).

DR EMBL: AP002998; BAB49309.1; -

DR InterPro: IPR001327; FAD_pyr_redox.

DR InterPro: IPR002938; MoxY_FAD binding.

DR InterPro: IPR000205; NAD_binding.

DR InterPro: IPR003042; Rng_moxxygenase.

DR Pfam: PF01494; FAD_binding_3; 1.

DR PRINTS: PR00368; FADPNR.

DR PRINTS: PR00420; RINGMOXGNASE.

KW Complete proteome.

SO SEQUENCE 406 AA; 44874 MW; 9A170BF328BDD2C0 CRC64;

Query Match 31.6%; Score 59; DB 16; Length 406;

Best local Similarity 39.4%; Pred. No. 38;

Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 GLRRLKFRNKIKKIKGKIQGLIPKLP 33

Db 371 GLEWRVRLNNARQLMTIGRKRHLGPRRP 403

RESULT 15

089297 PRELIMINARY: PRT: 400 AA.

AC 089297: PRELIMINARY: PRT: 400 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE POLYPROTEIN (FRAGMENT).

OS Yellow fever virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Flavivirus.

OX NCBI_TaxID=11089;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CAR77(883);

RX MEDLINE=9707878; PubMed=8918913;

RA Wang E., Weaver S.C., Shope R.E., Tesh R.B., Watts D.M., Barrett A.D.;

RT "Genetic variation in yellow fever virus: duplication in the 3'

RT noncoding region of strains from Africa";

RL Virology 225:274-281(1996).

DR EMBL: U52392; AAB01955.1; -

DR InterPro: IPR001122; Flav1_Capsid.

DR InterPro: IPR000336; Flav1_glycoprote.

DR InterPro: IPR000069; Flav1_M.

DR InterPro: IPR002535; Flav1_propep.

DR Pfam: PF01003; Flav1_capsid; 1.

DR Pfam: PF00869; Flav1_glycoprot; 1.

DR Pfam: PF01004; Flav1_M; 1.

DR Pfam: PF01570; Flav1_propep; 1.

DR ProDom: PD001556; Flav1_glycoprotef; 1.

FT NON_TER 400

SO SEQUENCE 400 AA; 44491 MW; 69191AA0FE674948 CRC64;

Query Match 31.0%; Score 58; DB 12; Length 400;

Best local Similarity 36.4%; Pred. No. 48;

Matches 12; Conservative 8; Mismatches 7; Indels 6; Gaps 1;

QY 2 LRRRLKFRNKIKKIKGKIQ-----IOGL 28

Db 16 VRRGVRSLSNKKIKGKIQIGNRPGPSRGVGF 48

Search completed: July 12, 2002, 08:03:51

Job time: 637 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:00:43 ; Search time 227.32 Seconds
(without alignments)
8.795 Million cell updates/sec

Title: US-09-642-744B-19
Perfect score: 89
Sequence: 1 KRLRKFRNKIKELKKIG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	18	22	AA870667
2	89	100.0	21	22	AA870666
3	89	100.0	22	22	AA870668
4	89	100.0	29	15	AA845671
5	89	100.0	29	21	AA807903
6	89	100.0	32	21	AA828486
7	89	100.0	37	22	AA870665
8	89	100.0	38	22	AA851193
9	79	88.8	19	22	AA870669
10	79	88.8	21	22	AA870670
11	75	84.3	15	22	AA870671

12	65	73.0	16	15	AA845673	RNTP heparin bindi
13	65	73.0	16	21	AA807905	Heparin-binding se
14	50	56.2	33	17	AA894449	Mouse antimicrobia
15	50	56.2	36	22	AA866421	Mouse cathelin pep
16	50	56.2	178	22	AA894612	Human protein sequ
17	49	55.1	18	20	AA857472	Antimicrobial pept
18	49	55.1	214	21	AA830858	Arabidopsis thalia
19	49	55.1	218	21	AA830857	Arabidopsis thalia
20	49	55.1	241	21	AA830856	Arabidopsis thalia
21	49	55.1	246	21	AA852394	Arabidopsis thalia
22	49	55.1	250	21	AA852393	Arabidopsis thalia
23	49	55.1	273	21	AA852392	Arabidopsis thalia
24	49	55.1	452	21	AA890302	1-2-rihamosyl-tren
25	48	53.9	15	22	AA870672	Lupine RCAP 18 cat
26	48	53.9	17	22	AA870673	Lupine RCAP 18 cat
27	48	53.9	439	22	AA834200	Staphylococcus aur
28	47	52.8	3421	22	AA884902	Chimeric protein f
29	46	51.7	18	13	AA821366	Sequence of amphip
30	46	51.7	18	13	AA822850	Amphiphilic peptid
31	46	51.7	123	19	AA875166	Human secreted pro
32	46	51.7	123	21	AA808924	Human secreted pro
33	46	51.7	123	21	AA808935	Human secreted pro
34	46	51.7	188	20	AA827071	Human JWA protein.
35	46	51.7	188	21	AA808900	Human secreted pro
36	46	51.7	188	22	AA893263	Human polypeptide,
37	46	51.7	188	22	AA801185	Rat glutamate tran
38	46	51.7	189	19	AA875110	Human secreted pro
39	46	51.7	157	21	AA829943	Arabidopsis thalia
40	45	50.6	174	21	AA829942	Arabidopsis thalia
41	45	50.6	183	21	AA829941	Arabidopsis thalia
42	45	50.6	269	21	AA857379	Bovine epimorphine
43	45	50.6	287	21	AA857378	Bovine epimorphine
44	45	50.6	287	21	AA857380	Sheep epimorphine
45	44	49.4	18	20	AA857503	Antimicrobial pept

ALIGNMENTS

RESULT 1
AAB70667 standard; peptide: 18 AA.
ID AAB70667:
AC AAB70667:
XX
XX
DT 15-MAY-2001 (first entry)
XX
XX
DE Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:20.
XX
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX
OS Oryctolagus cuniculus.
OS
XX
XX
PN WO200112668-A1.
XX
XX
PD 22-FEB-2001
XX
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
XX
PR 18-AUG-1999; 99US-0149886.
XX
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX
DR WPI: 2001-234911/24.
XX
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
and proliferation of microbes, and for treating microbial infections -


```
xx Claim 1; Page 103; 137pp; English.
PS
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
SQ Sequence 18 AA;

Query Match 100.0%; Score 89; DB 22; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.8e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRLRKFRNKIKKIKKIG 18
   |||||
DB 1 krlrkfrnkikiklkkig 18

RESULT 2
AAB70666
ID AAB70666 standard; peptide; 21 AA.
XX
XX AAB70666;
AC
XX
XX 15-MAY-2001 (first entry)
DT
XX
XX Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:19.
DE
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Oryctolagus cuniculus.
OS
XX
XX WO200112668-A1.
PN
XX
XX 22-FEB-2001.
PD
XX
XX 18-AUG-2000; 2000WO-US22781.
PF
XX
XX 18-AUG-1999; 99US-0149886.
PR
XX
XX (IOWA ) UNIV IOWA RES FOUND.
PA (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI
XX
XX WPI; 2001-234911/24.
DR
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
PT
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
```

```
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 89; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.5e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRLRKFRNKIKKIKKIG 18
   |||||
DB 4 krlrkfrnkikiklkkig 21

RESULT 3
AAB70668
ID AAB70668 standard; peptide; 22 AA.
XX
XX AAB70668;
AC
XX
XX 15-MAY-2001 (first entry)
DT
XX
XX Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:21.
DE
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Oryctolagus cuniculus.
OS
XX
XX WO200112668-A1.
PN
XX
XX 22-FEB-2001.
PD
XX
XX 18-AUG-2000; 2000WO-US22781.
PF
XX
XX 18-AUG-1999; 99US-0149886.
PR
XX
XX (IOWA ) UNIV IOWA RES FOUND.
PA (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
PI
XX
XX WPI; 2001-234911/24.
DR
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
PT
XX
XX Claim 1; Page 103; 137pp; English.
PS
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
XX
SQ Sequence 22 AA;

Query Match 100.0%; Score 89; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.8e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRLRKFRNKIKKIKKIG 18
   |||||
```

Db 2 krlrkfrnkikexlkkig 19

RESULT 4

AA045671 standard: Protein: 29 AA.

AA045671:

25-JUL-1994 (first entry)

N terminus of human CAP18.

Cationic antibacterial protein: lipopolysaccharide binding;

anticoagulant; granulocytes; RNIP; LPS; sepsis; autoimmune disorder;

septic shock; rabbit.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 23 /label= Asp, Lys

Misc-difference 26 /label= Gln, Ile

Misc-difference 27 /label= Gly, Gln

MO9402589-A.

03-FEB-1994.

15-JUL-1993; 93MO-US06731.

17-JUL-1992; 92US-0916761.

17-JUL-1992; 92US-0916765.

(PANO-) PANORAMA RES INC.

Hirata M, Larrick JW, Wright SC;

WPI: 1994-048847/06.

Sequences encoding mammalian cationic antibacterial proteins -

are homologous to human and rabbit CAP18 sequences and have

lipo:polysaccharide binding and anti-coagulation activity

Disclosure: Page 59; 112pp; English.

The sequence is that of a human cationic antibacterial protein CAP18

N-terminal fragment obtd. from granulocytes as sequenced. The

sequence corresponds to the N-terminus of the C-terminal RNIP fragment

of rabbit CAP18. The fragment is capable of binding to lipopolysaccharide

and inhibiting LPS-mediated activation of macrophage, as well as

interfering with the clotting cascade to inhibit coagulation in

conditions of disseminated intravascular coagulation. The polypeptides

can also be used to attenuate, inhibit or prevent LPS associated

conditions, e.g. sepsis, autoimmune disorders, inflammation, etc.

See also AA045667-81.

Sequence 29 AA;

Query Match 100.0%; Score 89; DB 15; Length 29;

Best Local Similarity 100.0%; Pred. No. 7.5e-06;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 krlrkfrnkikexlkkig 18

4 krlrkfrnkikexlkkig 21

RESULT 5

AA07903

ID AAB07903 standard: Protein: 29 AA.

AA07903;

14-NOV-2000 (first entry)

Putative N-terminal of a human cationic protein designated CAP18.

Human; cationic protein; lipopolysaccharide binding; anticoagulant;

CAP18; reactive nitrogen inhibitory protein; RNIP; gram negative sepsis;

coagulation-related disorder; disseminated intravascular coagulation.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 23 /label= Asp, Lys

Misc-difference 26 /label= Gln, Ile

Misc-difference 27 /label= Gly, Gln

US6103888-A.

15-AUG-2000.

01-JUN-1999; 99US-0322911.

01-AUG-1996; 96US-0691280.

17-JUL-1992; 92US-0916761.

17-JUL-1992; 92US-0916765.

15-JUL-1993; 93MO-US06731.

27-SEP-1994; 94US-0313681.

(PANO-) PANORAMA RES INC.

Larrick JW, Wright SC, Hirata M;

WPI: 2000-531989/48.

Novel human cDNA encoding cationic proteins having lipopolysaccharide

binding and anticoagulant activity, useful for treating and diagnosing

gram negative sepsis and disseminated intravascular coagulation -

Disclosure: Column 11; 46pp; English.

The present sequence represents the putative N-terminal of a human

cationic protein, having lipopolysaccharide binding and anticoagulant

activity. The polypeptide is designated CAP18. Amino acids 134-170

of CAP18 represent a reactive nitrogen inhibitory protein (RNIP).

The CAP18 polynucleotide is useful for producing cationic proteins.

The CAP18 polypeptide is useful for treating and diagnosing

lipopolysaccharide-associated conditions such as gram negative

sepsis, and/or coagulation-related disorders, such as disseminated

intravascular coagulation.

Sequence 29 AA;

Query Match 100.0%; Score 89; DB 21; Length 29;

Best Local Similarity 100.0%; Pred. No. 7.5e-06;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 krlrkfrnkikexlkkig 18

4 krlrkfrnkikexlkkig 21

RESULT 6

AA028486

AA028486 standard: peptide: 32 AA.

AA028486;

XX DR WPI; 2001-070965/08.
XX XX
PT Screening for nucleotide sequences encoding antimicrobial peptides by
PT generating peptide libraries in microorganisms, inducing expression of
PT peptides, selecting cells and recovering nucleotide sequences from
PT cells
XX
PS Example 1; Page 34; 59pp; English.
XX
CC The present invention describes a plasmid which is ligated with a pool
CC of nucleotide sequences (NT) linked to an inducible promoter, to express
CC a peptide (P) (an enzyme or mature (P) of less than 100 amino acids
CC optionally linked to a signal (P)), transformed into host cells and
CC cultured in presence of an inducer to induce expression of the NT. A
CC method of screening (1) a pool of nucleotide sequences to select a
CC nucleotide sequence encoding a peptide, comprises: (a) ligating a plasmid
CC with the pool of NT; (b) transforming host cells which are sensitive to
CC the peptide with the ligated plasmids; (c) screening the transformed
CC cells to select viable cells; (d) cultivating the viable cells in the
CC presence of an inducer to induce expression of NT; (e) selecting cells
CC according to the effect of the inducer on cell growth; and (f) recovering
CC NT encoding the peptide from the selected cells. (1) is useful for
CC screening a pool of nucleotide sequences to select a nucleotide sequence
CC encoding a peptide which is an antimicrobial peptide or an antimicrobial
CC enzyme active on bacteria and for finding and preparing a composition for
CC treatment of human or animal. The antimicrobial peptide obtained using
CC (1) may be employed in preservation of e.g. food/feed, paint
CC formulations, detergents, cosmetics, medical devices such as prosthetic
CC implants and also to disinfect and/or kill microbial cells on an object
CC e.g. as an disinfectant for the treatment of biofilm. The peptides are
CC useful for treating microbial infections and/or tumours. Peptides with
CC improved bio-activity can be developed using (1). The peptides have no
CC negative effect on normal mammalian and/or eukaryotic cells. The present
CC sequence represents an AMP (antimicrobial peptide) gene amino acid
CC sequence, which is used in an example from the present invention.
XX
SQ Sequence 38 AA:

Query Match 100.0%; Score 89; DB 22; Length 38;
Best Local Similarity 100.0%; Pred. No. 9,7e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRLKFRNKIKEXIKKIG 18
Db 5 KRLKFRNKIKEXIKKIG 22

RESULT 9
AAB70669
ID AAB70669 standard; peptide: 19 AA.
XX
AC AAB70669;
XX
DT 15-MAY-2001 (first entry)
XX
DE Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:22.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Oryctolagus cuniculus.
XX
XX WO200112668-A1.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
XX
PR

XX PA (IOWA) UNIV IOWA RES FOUND.
XX PA (REGC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
XX
PS Claim 1; Page 103; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (1)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
SQ Sequence 19 AA:

Query Match 88.8%; Score 79; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LKRFNRKIKEXIKKIG 18
Db 1 LKRFNRKIKEXIKKIG 16

RESULT 10
AAB70670
ID AAB70670 standard; peptide: 21 AA.
XX
AC AAB70670;
XX
DT 15-MAY-2001 (first entry)
XX
XX Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:23.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Oryctolagus cuniculus.
XX
XX WO200112668-A1.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
XX
XX (IOWA) UNIV IOWA RES FOUND.
XX (REGC) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections
XX
XX Claim 1; Page 103; 137pp; English.
XX
PS

XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX
SQ Sequence 21 AA;

Query Match 88.8%; Score 79; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LRKFRKKIKKKLKIG 18
Db 1 LRKFRKKIKKKLKIG 16

RESULT 11
AAB70671
ID AAB70671 standard; peptide; 15 AA.
XX
AC AAB70671;
XX
DT 15-MAY-2001 (first entry)
XX
DE Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:24.
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KM bactericidal; antiviral; microbial growth inhibitor;
KM proliferation inhibitor; gene therapy; *Pseudomonas aeruginosa*;
KM *Burkholderia cepacia*; *Alcaligenes*; *Xanthomonas*.
XX
XX *Oryctolagus cuniculus*.
OS
XX
PN WO200112668-A1.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
DR
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
XX
PS Claim 1: Page 103; 137pp; English.
XX
CC AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC antibiotic, antimicrobial and antiviral activities, and can be used as
CC microbial growth and proliferation inhibitors and in gene therapy. (I)
CC are useful for inhibiting microbial growth in an environment capable of
CC sustaining such growth, for inhibiting microbial growth or strain in a
CC host, and inhibiting the growth of drug-resistant microbial strains such
CC as *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Alcaligenes* and
CC *Xanthomonas*.
XX

XX
SQ Sequence 15 AA;

Query Match 84.3%; Score 75; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RKRFRKKIKKKLKIG 18
Db 1 RKRFRKKIKKKLKIG 15

RESULT 12
AAR45673
ID AAR45673 standard; Protein; 16 AA.
XX
AC AAR45673;
XX
DT 25-JUL-1994 (first entry)
XX
DE RNP heparin binding fragment.
XX
XX Cationic antibacterial protein; lipopolysaccharide binding;
KM anticoagulant; granulocytes; RNP; LPS; sepsis; autoimmune disorder;
KM septic shock; rabbit; CAP18.
XX
XX Synthetic.
OS
XX
PN WO9402589-A.
XX
PD 03-FEB-1994.
XX
PE 15-JUL-1993; 93WO-US06731.
XX
PR 17-JUL-1992; 92US-0916761.
PR 17-JUL-1992; 92US-0916765.
XX
PA (PANO-) PANORAMA RES INC.
XX
XX Hirata M, Larrick JW, Wright SC;
PI
XX
DR WPI; 1994-048847/06.
XX
PT Sequences encoding mammalian cationic antibacterial proteins -
PT are homologous to human and rabbit CAP18 sequences and have
PT lipo-polysaccharide binding and anti-coagulation activity
XX
PS Disclosure; Page 50; 112pp; English.
XX
XX The sequence of CAP18 C-terminal RNP was compared to that of a
CC number of heparin binding proteins to determine residues important for
CC binding to lipopolysaccharides and inhibiting LPS-mediated activation
CC of macrophage, as well as interfering with the clotting cascade to
CC inhibit coagulation in conditions of disseminated intravascular
CC coagulation.
CC See also AAR45667-81.
CC
XX
SQ Sequence 16 AA;

Query Match 73.0%; Score 65; DB 15; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0095;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRLKFRKKIKKK 13
Db 4 KRLKFRKKIKKK 16

RESULT 13
AAB07905
ID AAB07905 standard; Peptide; 16 AA.

```

XX AAB07905;
AC
XX
XX 14-NOV-2000 (first entry)
DT
XX
XX Heparin-binding sequence from a human cationic protein CAP18.
DE
XX
XX Human; cationic protein; lipopolysaccharide binding; anticoagulant;
KM CAP18; reactive nitrogen inhibitory protein; RNIP; gram negative sepsis;
KM coagulation-related disorder; disseminated intravascular coagulation;
KM heparin-binding.
XX
XX Homo sapiens.
OS
XX
XX US6103888-A.
PN
XX
XX 15-AUG-2000.
PD
XX
XX 01-JUN-1999; 99US-0322911.
PF
XX
XX 01-AUG-1996; 96US-0691280.
PR 17-JUL-1992; 92US-0916761.
PR 17-JUL-1992; 92US-0916765.
PR 15-JUL-1993; 93WO-US06731.
PR 27-SEP-1994; 94US-0313681.
XX
XX (PANO-) PANORAMA RES INC.
PA
XX
XX Larrick JM, Wright SC, Hirata M;
PI
XX
XX WPI; 2000-531989/48.
DR
XX
XX Novel human cDNA encoding cationic proteins having lipopolysaccharide
PT binding and anticoagulant activity; useful for treating and diagnosing
PT gram negative sepsis and disseminated intravascular coagulation -
PT
XX
XX Disclosure; Column 26; 46pp; English.
PS
XX
XX The present sequence represents a heparin-binding sequences, derived
CC from a human cationic protein, having lipopolysaccharide binding and
CC anticoagulant activity. The polypeptide is designated CAP18. Amino
CC acids 134-170 of CAP18 represent a reactive nitrogen inhibitory
CC protein (RNIP). The CAP18 polynucleotide is useful for producing
CC cationic proteins. The CAP18 polypeptide is useful for treating and
CC diagnosing lipopolysaccharide-associated conditions such as gram
CC negative sepsis, and/or coagulation-related disorders,
CC such as disseminated intravascular coagulation.
CC
XX
XX Sequence 16 AA:
SQ

```

```

Query Match 73.0%; Score 65; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0095;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KRLRKFRNKIKK 13
   ||| ||| ||| |||
DB 4 krlrkfrnkik 16

```

```

RESULT 14
AAR94449
ID AAR94449 standard; peptide; 33 AA.
XX
XX AAR94449;
AC
XX
XX 05-NOV-1996 (first entry)
DT
XX
XX Mouse antimicrobial peptide ineffective syndecan expression inducer.
DE
XX
XX Induction; expression; syndecan-1; syndecan-4; surface;
KM mesenchymal cell; fibroblast; epithelial; ineffective.
XX
XX

```

```

OS Mus musculus.
XX
XX W09609322-A2.
PN
XX
XX 28-MAR-1996.
PD
XX
XX 22-SEP-1995; 95WO-US12080.
PF
XX
XX 22-SEP-1994; 94US-0310722.
PR
XX
XX (CHIL-) CHILDRENS MEDICAL CENT.
PA
XX
XX Bernfield M, Gallo RL;
PI
XX
XX WPI; 1996-188401/19.
DR
XX
XX Modulating mesenchymal interaction by administration of syndecan
PT used in the treatment of wounds, tumours, restenosis, etc
PT
XX
XX Example 8; Page 27; 34pp; English.
PS
XX
XX The present peptide, a mouse antibacterial peptide, is an
CC ineffective syndecan-1 and syndecan-4 mesenchymal cell, esp.
CC fibroblast and epithelial cell surface expression inducer.
CC Human microvascular endothelial cells were assayed for syndecan-4
CC expression following exposure to 5 % wound fluid, dbcAMP (1 mM),
CC the present peptide (10 microm) or a blank, to give respective
CC cell surface syndecan-4 values (MOD/m in) of approx. 1.75, 1.70,
CC 1.00 and 0.95.
XX
XX
XX Sequence 33 AA:
SQ

```

```

Query Match 56.2%; Score 50; DB 17; Length 33;
Best Local Similarity 75.0%; Pred. No. 2.4;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 3 LKFRNKIKKIKKIG 18
   ||| ||| ||| ||| |||
DB 3 lrxggekigekikkg 18

```

```

RESULT 15
AAG66421
ID AAG66421 standard; peptide; 36 AA.
XX
XX
XX AAG66421;
AC
XX
XX 23-OCT-2001 (first entry)
DT
XX
XX Mouse cathelin peptide.
DE
XX
XX Mouse; cathelin; immunomodulator; polycationic polymer; vaccine;
KM immune response.
KM
XX
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
FH
FT MISC-difference 22
FT /note="Residue given as "O" in specification"
XX
XX
XX W0200154720-A1.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 05-JAN-2001; 2001WO-EP00087.
PF
XX
XX 28-JAN-2000; 2000AT-0000129.
PR
XX
XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
PA
XX
XX Lingnau K, Maltner F, Schmidt W, Birnstiel M, Buschle M;
PI
XX
XX

```

DR WPI, 2001-536419/59.
XX
PT Pharmaceutical composition useful for inducing immune response
PT comprises antigen, immunogenic oligodeoxynucleotide containing
PT cytosine-guanine dinucleotide motifs and polycationic polymer
XX
PS Disclosure; Page 12; 39pp; English.
XX
CC The present invention relates to a pharmaceutical composition which
CC comprises an antigen, an immunogenic oligodeoxynucleotide containing
CC cytosine-guanine dinucleotide (CpG) motifs (CpG-ODN) and a polycationic
CC polymer. The composition is useful for making a vaccine to induce potent
CC immune responses, or to decrease or ablate undesired immune responses.
CC The present sequence is mouse cathelin peptide, which was used as a
CC polycationic compound in the present invention.
XX
SQ Sequence 36 AA;

Query Match 56.2%; Score 50; DB 22; Length 36;
Best Local Similarity 75.0%; Pred. No. 2.6;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 LRRFRNKIKEKLKKG 18
||| || |||||
Db 6 lrxggekigeklkkig 21

Search completed: July 12, 2002, 08:00:44
Job time: 451 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:55:08 ; Search time 75.52 Seconds
(without alignments)
5.822 Million cell updates/sec

Title: US-09-642-744B-19
Perfect score: 89
Sequence: 1 KRLKFRNKIKELKIG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	29	1 US-08-313-681A-7	Sequence 7, Appl
2	89	100.0	29	3 US-09-322-911-7	Sequence 7, Appl
3	89	100.0	171	1 US-08-313-681A-4	Sequence 4, Appl
4	89	100.0	171	3 US-09-322-911-4	Sequence 4, Appl
5	65	73.0	16	1 US-08-313-681A-11	Sequence 11, Appl
6	65	73.0	16	3 US-09-322-911-11	Sequence 11, Appl
7	50	56.2	33	5 PCT-US95-12080-4	Sequence 4, Appl
8	46	51.7	18	1 US-07-725-331-29	Sequence 29, Appl
9	46	51.7	18	3 PCT-US91-05047-29	Sequence 29, Appl
10	44	49.4	187	3 US-08-493-071-16	Sequence 16, Appl
11	44	49.4	236	3 US-08-493-071-15	Sequence 15, Appl
12	44	49.4	277	1 US-08-690-457-5	Sequence 5, Appl
13	44	49.4	277	2 US-08-628-187-5	Sequence 5, Appl
14	44	49.4	277	3 US-08-628-187-3	Sequence 3, Appl
15	44	49.4	287	1 US-08-690-457-4	Sequence 4, Appl
16	44	49.4	287	2 US-08-628-187-4	Sequence 4, Appl
17	44	49.4	287	3 US-08-493-071-2	Sequence 2, Appl
18	44	49.4	288	1 US-08-690-457-3	Sequence 3, Appl
19	44	49.4	288	2 US-08-628-187-3	Sequence 3, Appl
20	44	49.4	288	3 US-08-493-071-1	Sequence 1, Appl
21	43	48.3	18	1 US-07-725-331-24	Sequence 24, Appl
22	43	48.3	18	3 PCT-US91-05047-24	Sequence 24, Appl
23	43	48.3	18	5 PCT-US91-05047-27	Sequence 27, Appl
24	43	48.3	18	5 PCT-US91-05047-27	Sequence 27, Appl
25	42	47.2	154	2 US-08-710-330A-11	Sequence 11, Appl
26	42	47.2	303	1 US-08-109-391A-2	Sequence 2, Appl
27	42	47.2	303	1 US-08-459-019A-2	Sequence 2, Appl

28	42	47.2	303	2 US-08-460-428A-2	Sequence 2, Appl
29	42	47.2	303	3 US-08-458-860A-2	Sequence 2, Appl
30	41	46.1	21	1 US-07-965-663A-3	Sequence 3, Appl
31	41	46.1	187	3 US-08-493-071-19	Sequence 19, Appl
32	41	46.1	236	3 US-08-493-071-18	Sequence 18, Appl
33	41	46.1	279	1 US-08-690-457-11	Sequence 11, Appl
34	41	46.1	279	2 US-08-628-187-11	Sequence 11, Appl
35	41	46.1	279	3 US-08-493-071-6	Sequence 6, Appl
36	41	46.1	288	1 US-08-690-457-10	Sequence 10, Appl
37	41	46.1	288	2 US-08-628-187-10	Sequence 10, Appl
38	41	46.1	288	3 US-08-493-071-5	Sequence 5, Appl
39	41	46.1	289	1 US-08-690-457-9	Sequence 9, Appl
40	41	46.1	289	2 US-08-628-187-9	Sequence 9, Appl
41	41	46.1	289	3 US-08-493-071-4	Sequence 4, Appl
42	41	46.1	343	2 US-08-599-171A-28	Sequence 28, Appl
43	41	46.1	343	3 US-08-646-580B-28	Sequence 28, Appl
44	41	46.1	343	3 US-09-069-226-28	Sequence 28, Appl
45	41	46.1	343	4 US-09-412-184-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-08-313-681A-7
Sequence 7, Application US/08313681A
Patent No. 5618675
GENERAL INFORMATION:
APPLICANT: Larrick, James W.
APPLICANT: Wright, Susan C.
APPLICANT: Hirata, Mishimasa
TITLE OF INVENTION: Human Cationic Proteins Having
TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESSES:
ADDRESS: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,681A
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15325-9-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 23
OTHER INFORMATION: /note= "Xaa is Asp or Lys"
FEATURE:
NAME/KEY: Region
LOCATION: 26
OTHER INFORMATION: /note= "Xaa is a Gln or Ile"
FEATURE:

NAME/KEY: Region
LOCATION: 27
OTHER INFORMATION: /note= "Xaa is a Gly or Gln"
US-08-313-681A-7

Query Match 100.0%; Score 89; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRLKFRNKIKELKKG 18
|||||
Db 4 KRLKFRNKIKELKKG 21

RESULT 2
US-09-322-911-7
Sequence 7, Application US/09322911
Patent No. 6103888
GENERAL INFORMATION:
APPLICANT: Larrick, James W.
APPLICANT: Wright, Susan C.
APPLICANT: Hirata, Mishimasa
APPLICANT: Balint, Robert F.
TITLE OF INVENTION: Human Cationic Proteins Having
TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/322,911
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/691,280
FILING DATE: August 1, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,761
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,765
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06731
FILING DATE: July 15, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/313,681
FILING DATE: September 27, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 15325-000920
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:

NAME/KEY: Region
LOCATION: 23
OTHER INFORMATION: /note= "Xaa is Asp or Lys"
FEATURE:
NAME/KEY: Region
LOCATION: 26
OTHER INFORMATION: /note= "Xaa is a Gln or Ile"
FEATURE:
NAME/KEY: Region
LOCATION: 27
OTHER INFORMATION: /note= "Xaa is a Gly or Gln"
US-09-322-911-7

Query Match 100.0%; Score 89; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRLKFRNKIKELKKG 18
|||||
Db 4 KRLKFRNKIKELKKG 21

RESULT 3
US-08-313-681A-4
Sequence 4, Application US/08313681A
Patent No. 5618675
GENERAL INFORMATION:
APPLICANT: Larrick, James W.
APPLICANT: Wright, Susan C.
APPLICANT: Hirata, Mishimasa
APPLICANT: Balint, Robert F.
TITLE OF INVENTION: Human Cationic Proteins Having
TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,681A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541
REFERENCE/DOCKET NUMBER: 15325-9-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 171 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-313-681A-4
Query Match 100.0%; Score 89; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KRLKFRNKIKELKKG 18
|||||

Db 138 KRLKFRNKIKELKKG 155

RESULT 4

US-09-322-911-4
; Sequence 4, Application US/09322911
; Patent No. 6103888
; GENERAL INFORMATION:
; APPLICANT: Larick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishima
; APPLICANT: Balint, Robert F.
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/322,911
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/691,280
; FILING DATE: August 1, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,761
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,765
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06731
; FILING DATE: July 15, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/313,681
; FILING DATE: September 27, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 15325-000920
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-322-911-4

Query Match 100.0%; Score 89; DB 3; Length 171;

Best Local Similarity 100.0%; Pred. No. 5; Le-05; Mismatches 0; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 138 KRLKFRNKIKELKKG 155

RESULT 5

US-08-313-681A-11

; Sequence 11, Application US/08313681A
; Patent No. 5618675
; GENERAL INFORMATION:
; APPLICANT: Larick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishima
; APPLICANT: Balint, Robert F.
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,681A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15325-9-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-313-681A-11

Query Match 73.0%; Score 65; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.008; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 4 KRLKFRNKIKELKKG 16

US-09-322-911-11

; Sequence 11, Application US/09322911
; Patent No. 6103888
; GENERAL INFORMATION:
; APPLICANT: Larick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishima
; APPLICANT: Balint, Robert F.
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/322,911
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/691,280
FILING DATE: August 1, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,761
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,765
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06731
FILING DATE: July 15, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/313,681
FILING DATE: September 27, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 15325-000920
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-322-911-11

Query Match      73.0%; Score 65; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.008;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKFRNKIKK 13
    |111111111111
Db 4 KRLKFRNKIKK 16

RESULT 7
PCT-US95-12080-4
Sequence 4, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporation
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-12080-4

Query Match      56.2%; Score 50; DB 5; Length 33;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LKFRNKIKKIKKIG 18
    |111111111111
Db 3 LKRGKIKGKIKKIG 18

RESULT 8
US-07-725-331-29
Sequence 29, Application US/07725331
Patent No. 5294605
GENERAL INFORMATION:
APPLICANT: Houghten, Richard
TITLE OF INVENTION: Amphiphilic Peptide Compositions and
ANALOGUES THEREOF
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSER: Dressler, Goldsmith, Sutker, Shore,
STREET: 180 No. 5294605th Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/725,331
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/554,422
FILING DATE: 19-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Gansson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: 421250-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3126165418
TELEFAX: 3126165460
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
OTHER INFORMATION: May be a C-terminal amide, and/or may
be acetylated at N-terminus.
US-07-725-331-29

Query Match      51.7%; Score 46; DB 1; Length 18;
Best Local Similarity 52.9%; Pred. No. 2.7;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
```


TELEPHONE: 703-684-1111
TELEFAX: 703-684-1124
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-493-071-15

Query Match 49.4%; Score 44; DB 3; Length 236;
Best Local Similarity 52.9%; Pred. No. 52;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KRLKFRNKIKKELKKI 17
|::| |||: ||| |
Db 51 KEIKTKANKIRAKLKAI 67

RESULT 12
US-08-690-457-5
Sequence 5, Application US/08690457
Patent No. 5726298
GENERAL INFORMATION:
APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.
TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE
TITLE OF INVENTION: DESIGNATED AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANTIBOD
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA,
ADDRESS: Japan
STREET: 1, Taya-cho
City: Yokohama-shi
COUNTRY: Japan
ZIP: 244
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690.457
FILING DATE: 16-AUG-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/078,309
FILING DATE: June 15, 1993
APPLICATION NUMBER: 294856/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 294857/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 122906/1992
FILING DATE: April 17, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 135692/1992
FILING DATE: April 30, 1992
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 277
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-690-457-5

Query Match 49.4%; Score 44; DB 1; Length 277;
Best Local Similarity 52.9%; Pred. No. 60;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KRLKFRNKIKKELKKI 17
|::| |||: ||| |
Db 79 KEIKTKANKIRAKLKAI 95

RESULT 13
US-08-628-187-5
Sequence 5, Application US/08628187
Patent No. 5837239
GENERAL INFORMATION:
APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.
TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE DESIGNATED
TITLE OF INVENTION: AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANTIBODIES THE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA, Japan
STREET: 1, Taya-cho
City: Yokohama-shi
COUNTRY: Japan
ZIP: 244
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628.187
FILING DATE: April 5, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 294856/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 294857/1991
FILING DATE: October 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 122906/1992
FILING DATE: April 17, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 135692/1992
FILING DATE: April 30, 1992
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 277
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-628-187-5

Query Match 49.4%; Score 44; DB 2; Length 277;
Best Local Similarity 52.9%; Pred. No. 60;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KRLKFRNKIKKELKKI 17
|::| |||: ||| |
Db 79 KEIKTKANKIRAKLKAI 95

RESULT 14
US-08-493-071-3
Sequence 3, Application US/08493071
Patent No. 6127149
GENERAL INFORMATION:
APPLICANT: Hirai, Yohei
APPLICANT: Koshida, Shogo
APPLICANT: Oka, Yumiko
TITLE OF INVENTION: MODIFIED EPIMORPHIN
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
STREET: 99 CANAL CENTER PLAZA, SUITE 300

```

: CITY: ALEXANDRIA
: STATE: VA
: COUNTRY: USA
: ZIP: 22314
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/493,071
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Price, Robert L.
: REGISTRATION NUMBER: 22,685
: REFERENCE/DOCKET NUMBER: 715-107
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-684-1111
: TELEFAX: 703-684-1124
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 277 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-493-071-3

```

```

Query Match          49.4%; Score 44; DB 3; Length 277;
Best Local Similarity 52.9%; Pred. No. 60;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 KRLKFRNKIKKELKKI 17
   1 : : 1 111 : 111 1
Db 79 KEIKTKANKIRAKLKAI 95

```

```

RESULT 15
US-08-690-457-4
: Sequence 4, Application US/08690457
: Patent No. 5726298
: GENERAL INFORMATION:
: APPLICANT: BIOMATERIAL RESEARCH INSTITUTE CO., LTD.
: TITLE OF INVENTION: NOVEL PHYSIOLOGICALLY ACTIVE SUBSTANCE
: TITLE OF INVENTION: DESIGNATED AS EPIMORPHIN, GENES ENCODING THE SAME, AND ANTIBOD
: NUMBER OF INVENTION: 15
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: 1, Taya-cho, Sakae-ku, Yokohama-shi, KANAGAWA,
: STREET: 1, Taya-cho
: CITY: Yokohama-shi
: COUNTRY: Japan
: ZIP: 244
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/690,457
: FILING DATE: 16-AUG-1996
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/078,309
: FILING DATE: June 15, 1993
: APPLICATION NUMBER: 294856/1991
: FILING DATE: October 16, 1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 294857/1991
: FILING DATE: October 16, 1991

```

```

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 122906/1992
: FILING DATE: April 17, 1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 135692/1992
: FILING DATE: April 30, 1992
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 287
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-690-457-4

```

```

Query Match          49.4%; Score 44; DB 1; Length 287;
Best Local Similarity 52.9%; Pred. No. 62;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 KRLKFRNKIKKELKKI 17
   1 : : 1 111 : 111 1
Db 79 KEIKTKANKIRAKLKAI 95

```

```

Search completed: July 12, 2002, 07:55:09
Job time: 116 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:56:55 ; Search time 95.47 Seconds

(Without alignments)
18.117 Million cell updates/sec

Title: US-09-642-744B-19

Perfect score: 89

Sequence: 1 KRLKFRNKIKELKIKIG 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	100.0	171	2	18K lipopolysaccha
2	51	57.3	241	2	hypothetical prote
3	50	56.2	166	2	antibacterial prot
4	50	56.2	173	2	cathelin-related p
5	49	55.1	142	2	hypothetical prote
6	49	55.1	324	2	probable choline-p
7	49	55.1	326	2	choline-phosphate
8	48	53.9	497	2	hypothetical prote
9	47.5	53.4	274	2	conserved hypotet
10	47	52.8	121	2	conserved hypotet
11	47	52.8	259	2	14-3-3 brain prote
12	47	52.8	1163	1	genome polypeptid
13	47	52.8	3411	1	genome polypeptid
14	47	52.8	3411	1	genome polypeptid
15	46	51.7	169	2	polypeptide deform
16	46	51.7	305	2	hypothetical prote
17	46	51.7	329	2	probable choline-p
18	46	51.7	331	2	probable choline-p
19	46	51.7	444	2	hypothetical prote
20	46	51.7	984	2	hypothetical prote
21	46	51.7	996	1	preproteins transio
22	45.5	51.1	661	2	NAD+ Adb-ribosyltr
23	45	50.6	153	2	hypothetical prote
24	45	50.6	174	2	myoglobin - Baikal
25	45	50.6	463	2	hypothetical prote
26	45	50.6	522	2	probable seryl-trn
27	45	50.6	522	2	hypothetical prote
28	45	50.6	522	2	type I site-specif
29	45	50.6	655	2	hypothetical prote
30	45	50.6	831	2	K06H7.3 protein -

30	45	50.6	1233	2	154383	chromosome segrega
31	44.5	50.0	458	2	D70410	cytosolic axial fi
32	44	49.4	96	2	H69042	ribosomal protein
33	44	49.4	109	2	G64379	hypothetical prote
34	44	49.4	184	2	T04396	Rich5 protein - ba
35	44	49.4	254	2	H86355	probable 14-3-3 pr
36	44	49.4	298	2	JN0466	epimorphin - human
37	44	49.4	298	2	E85166	probable phosphoch
38	44	49.4	425	2	T24111	hypothetical prote
39	44	49.4	894	2	T15769	hypothetical prote
40	43.5	48.9	477	2	A75052	cysteiny1-trNA syn
41	43.5	48.9	573	2	D90202	methionyl-trNA syn
42	43.5	48.9	996	2	B84481	hypothetical prote
43	43	48.3	73	2	T03182	hypothetical prote
44	43	48.3	96	2	F72228	ribosomal protein
45	43	48.3	119	2	C90351	hypothetical prote

ALIGNMENTS

```

RESULT 1
J01171
18K lipopolysaccharide-binding protein precursor - rabbit
N:Alternate names: 18K cationic protein
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 06-Dec-1996
C:Accession: J01171; PS0226
R:Larriek, J.W.; Morgan, J.G.; Palings, I.; Hirta, M.; Yen, M.H.
Biochem. Biophys. Res. Commun. 179, 170-175, 1991
A:Title: Complementary DNA sequence of rabbit CAP18-a unique lipopolysaccharide bindi
A:Reference number: J01171; MUID:91354246
A:Accession: J01171
A:Molecule type: mRNA
A:Residues: 1-171 <LAR>
A:Experimental source: Done marrow
A:Accession: PS0226
A:Molecule type: Protein
A:Residues: 135-159, 'OIGOL' <LA2>
A>Note: 157-Asp was also found
C:Superfamily: cathelin: cystatin homology
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-171/Product: 18K lipopolysaccharide-binding protein #status predicted <MAT>

Query Match 100.0%; Score 89; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. NO. 7.5e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRLKFRNKIKELKIKIG 18
Db 138 KRLKFRNKIKELKIKIG 155

RESULT 2
G86355
hypothetical protein T16E15.9 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: G86355
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucero, J.S.; Matli, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: G86355
A>Status: preliminary

```



```

genome polyprotein - yellow fever virus (strain 1899/81) (fragment)
N:Contains: amino end of nonstructural protein NS1; capsid protein C; envelope protein M
C:Species: yellow fever virus
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001
C:Accession: J00374
R:Ballinger-Crabtree, M.E.; Miller, B.R.
J. Gen. Virol. 71, 2115-2121, 1990
A:Title: Partial nucleotide sequence of South American yellow fever virus strain 1899/81
A:Reference number: J00374; M01D:91011358
A:Accession: J00374
A:Molecule type: genomic RNA
A:Cross-references: GB:D14458; GB:D00739; NID:9222777; PIDN:BAA03355.1; PID:9222778
C:Superfamily: yellow fever virus genome polyprotein
C:Keywords: capsid protein; envelope protein; glycoprotein; nonstructural protein; nucle
F:2-11/Domain: capsid protein C #status predicted <TM4>
F:106-122/Domain: transmembrane #status predicted <TM1>
F:122-285/Domain: transmembrane #status predicted <TM2>
F:251-267/Domain: transmembrane #status predicted <TM3>
F:271-287/Domain: transmembrane #status predicted <TM4>
F:286-778/Domain: major envelope protein E #status predicted <ENP>
F:733-753/Domain: transmembrane #status predicted <TM4>
F:762-778/Domain: transmembrane #status predicted <TM5>
F:779-1163/Domain: nonstructural protein NS1 (fragment) #status predicted <NS1>
F:1068-1075/Region: nucleotide-binding motif A (P-loop)
F:1133-1151/Domain: transmembrane #status predicted <TM6>
F:134,150,172,266,554,594,755,908,986/Binding site: carbohydrate (Asn) (covalent) #statu

```

```

F:2507-3411/Product: nonstructural protein NS5 #status predicted <NS5>
F:134,150,172,266,554,755,908,986,1756,2062,2320,2346,2408,2467,2720,2734,2740/Bindin
QY 3 LRKFRNKIKKIKKIG 18
: | ||||: | |||
Db 20 VRSLNKKIKOKTKOIG 35

Query Match 52.8%; Score 47; DB 1; Length 3411;
Best Local Similarity 56.2%; Pred. No. 4.3e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

genome polyprotein - yellow fever virus (strain Pasteur 17D-204)
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstruct
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: yellow fever virus
C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 19-Jan-2001
C:Accession: S07757
R:Dupuy, A.; Despres, P.; Cahour, A.; Girard, M.; Bouloy, M.
Nucleic Acids Res. 17, 3989, 1989
A:Title: Nucleotide sequence comparison of the genome of two 17D-204 yellow fever vac
A:Reference number: S07757; M01D:89282413
A:Accession: S07757
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-3411 <DUP>
A:Cross-references: EMBL:X15062; NID:962289; PIDN:CA837419.1; PID:94456986
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1989, i
C:Superfamily: yellow fever virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protei
F:2-210/Product: capsid protein C #status predicted <CPC>
F:105-125/Domain: transmembrane #status predicted <TM1>
F:211-285/Product: envelope protein M #status predicted <ENP>
F:271-289/Domain: transmembrane #status predicted <TM2>
F:286-778/Product: major envelope protein E #status predicted <ME>
F:736-753/Domain: transmembrane #status predicted <TM3>
F:756-778/Domain: transmembrane #status predicted <TM4>
F:779-1187/Domain: nonstructural protein NS1 #status predicted <NS1>
F:1133-1151/Domain: transmembrane #status predicted <TM5>
F:1160-1179/Domain: transmembrane #status predicted <TM6>
F:1188-1354/Product: nonstructural protein NS2a #status predicted <NS2a>
F:1355-1484/Product: nonstructural protein NS2b #status predicted <NS2b>
F:1485-2107/Product: nonstructural protein NS3 #status predicted <NS3>
F:1682-1689/Region: nucleotide-binding motif A (P-loop)
F:1769-1774/Region: nucleotide-binding motif B
F:1773-1776/Region: DEAH motif
F:2108-2394/Product: nonstructural protein NS4a #status predicted <NS4a>
F:2395-2506/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2507-3411/Product: nonstructural protein NS5 #status predicted <NS5>
F:134,150,172,594,908,986,1796,2062,2320,2346,2408,2467,2720,2734,2740/Binding site:

```

Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A:Reference number: A82035; MUID:20406833
 A:Accession: A82373
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-169 <HEI>
 A:Cross-References: GB:AE004096; GB:AE003852; NID:9954440; PIDN:AAF93224.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC0046
 A:Map position: 1
 C:Superfamily: polypeptide deformylase

Query Match 51.7%; Score 46; DB 2; Length 169;
 Best Local Similarity 81.8%; Pred. No. 37;
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 RNKIKKELKKI 17
 ||:|||||:
 DB 152 RNRIKELKKI 162

Search completed: July 12, 2002, 07:56:57
 Job time: 224 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:04:41 ; Search time 46.15 Seconds

(Without alignments)
15.102 Million cell updates/sec

Title: US-09-642-744B-19

Perfect score: 89
Sequence: 1 KRLRRFRNKIKELKKIG 18

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt-40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match length	ID	Description
1	89	100.0	171	CP18_RABIT	P25230 oryctolagus
2	50	56.2	166	MP36_PIG	P49931 sus scrofa
3	50	56.2	173	CRAM_MOUSE	P51437 mus musculus
4	49	55.1	241	143B_ARATH	O95928 arabidopsis
5	47	52.8	259	1433_CHLRE	P52908 chlamydomon
6	47	52.8	163	POLG_YEPV8	P29165 yellow fever
7	47	52.8	3411	POLG_YEPV1	P03314 y genome po
8	47	52.8	3411	POLG_YEPV2	P19901 y genome po
9	46	51.7	169	DEF_VIRCH	O9KXV3 vibrio chol
10	46	51.7	984	SECA_AQUAE	O67718 aquifex aeo
11	46	51.7	996	PPOL_SARPE	O11208 sarcophaga
12	46	51.7	1111	YJEP_HAEIN	O57362 haemophilus
13	45	50.6	153	MYG_PHOSI	P30562 phoca sibir
14	45	50.6	463	SYS_AERPE	O9YA63 aeropyrum p
15	45	50.6	618	YMX3_CABEL	P34511 caenorhabdi
16	45	50.6	655	YJ68_YEAST	P47139 saccharomyc
17	44	49.4	96	RL21_METTH	O27378 methanobact
18	44	49.4	109	Y639_METJA	O58056 methanococ
19	44	49.4	254	143A_ARATH	P48347 arabidopsis
20	43.5	48.9	477	SYC_PIRAB	O9UYV2 pyrococcus
21	43	48.3	96	RS20_THEMA	O9J177 thermotoga
22	43	48.3	155	YF45_PSEAE	O04628 pseudomonas
23	43	48.3	176	YL57_ARCFU	O28125 archaeoglob
24	43	48.3	252	Y830_METJA	O58240 methanococ
25	43	48.3	356	Y670_METJA	O58084 methanococ
26	43	48.3	476	SYC_PIRAB	O58370 pyrococcus
27	43	48.3	525	YI07_METJA	O67501 methanococ
28	43	48.3	555	YLVD_AQUAE	O67009 aquifex aeo
29	43	48.3	1630	MSPI_PLAFK	P04932 plasmodium
30	43	48.3	1639	MSPI_PLAFV	P04933 plasmodium
31	42.5	47.8	308	Y227_METJA	O57660 methanococ
32	42	47.2	153	MYG_CANFA	P02158 canis fami1
33	42	47.2	153	MYG_CTBGU	P20856 ctenodactyl

34	42	47.2	153	1	MYG_ELENA	P02186 elephas max
35	42	47.2	153	1	MYG_ERIEU	P02156 erinaceus e
36	42	47.2	153	1	MYG_GALCR	P02168 galago ctes
37	42	47.2	153	1	MYG_HALCR	P02162 halichoerus
38	42	47.2	153	1	MYG_HORSE	P02188 equus caball
39	42	47.2	153	1	MYG_LAGMA	P04250 lagostomus
40	42	47.2	153	1	MYG_LEPMU	P02169 lepidemur m
41	42	47.2	153	1	MYG_LOXAF	P02187 loxodonta a
42	42	47.2	153	1	MYG_LYCP1	P02159 lycodon pict
43	42	47.2	153	1	MYG_MEIME	P02157 meles meles
44	42	47.2	153	1	MYG_MOUSE	P04247 mus musculu
45	42	47.2	153	1	MYG_NYCCO	P02167 nycticebus

ALIGNMENTS

RESULT 1
CP18_RABIT STANDARD; PRT; 171 AA.
AC P25230;
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Antimicrobial protein CAP18 precursor (18 kDa lipopolysaccharide-
binding protein) (18 kDa cationic protein) (CAP18-A).
GN CAP18
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 135-159.
RC TISSUE=Bone marrow;
RX MEDLINE=91354246; PubMed=1883348;
RA Larrick J.W., Morgan J.G., Pallings I., Hirata M., Yen M.H.;
RT "Complementary DNA sequence of rabbit CAP18 -- a unique
lipopolysaccharide binding protein.";
RL Biochem. Biophys. Res. Commun. 179:170-175(1991).
RN [2]
RP SEQUENCE OF 135-159, AND CHARACTERIZATION.
RX MEDLINE=94178952; PubMed=8132348;
RA Hirata M., Shimomura Y., Yoshida M., Morgan J.G., Pallings I.,
RA Wilson D., Yen M.H., Wright S.C., Larrick J.W.;
RT "Characterization of a rabbit cationic protein (CAP18) with
lipopolysaccharide-inhibitory activity.";
RL Infect. Immun. 62:1421-1426(1994).
RN [3]
RP SEQUENCE OF 135-154, AND CHARACTERIZATION.
RX MEDLINE=94075827; PubMed=8254193;
RA Larrick J.W., Hirata M., Zheng H., Zhong J., Bolin D.,
RA Cavallion J.-M., Warren R.S., Wright S.C.;
RT "A novel granulocyte-derived peptide with lipopolysaccharide-
neutralizing activity.";
RL J. Immunol. 152:231-240(1994).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=94148064; PubMed=8313956;
RA Tossi A., Scocchi M., Skerlavaj B., Gennaro R.;
RT "Identification and characterization of a primary antibacterial
domain in CAP18, a lipopolysaccharide binding protein from rabbit
leukocytes.";
RL FEBS Lett. 339:108-112(1994).
RN [5]
RP STRUCTURE BY NMR OF 135-166.
RX MEDLINE=95377455; PubMed=7649303;
RA Chen C., Brock R., Luh F., Chou P.-J., Larrick J.W., Huang R.-F.,
RA Huang T.-H.;
RT "The solution structure of the active domain of CAP18 -- a
lipopolysaccharide binding protein from rabbit leukocytes.";
RL FEBS Lett. 370:46-52(1995).
CC -!- FUNCTION: CAP18 BINDS TO THE LIPID A MOIETY OF BACTERIAL
LIPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER

```
CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. HAS ANTIBIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: NEUTROPHILS.
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M73998; AAA3187.1; -.
CC PIR: J01171; J01171.
CC PDB: 1LTP; 31-MAR-95.
CC InterPro: IPR001894; Cathelicidin.
CC Pfam: PF00666; Cathelicidins; 1.
CC ProDom: PD001838; Cathelicidins; 1.
CC PROSITE: PS00946; CATHELICIDINS_1; 1.
CC PROSITE: PS00947; CATHELICIDINS_2; 1.
CC Antibiotic; Signal; 3D-structure.
CC FT SIGNAL 1 29 POTENTIAL.
CC FT CHAIN 30 171 ANTIMICROBIAL PROTEIN CAP18.
CC FT CHAIN 135 171 ANTIMICROBIAL PROTEIN CAP7.
CC FT MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID
CC (BY SIMILARITY).
CC FT DISULFID 85 96 BY SIMILARITY.
CC FT DISULFID 107 124 BY SIMILARITY.
CC FT VARIANT 157 157 K -> D.
CC SEQUENCE 171 AA; 19805 MW; D7BF2103BCF813C4 CRC64;

Query Match 100.0%; Score 89; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRLRKFRNKIKKKIKG 18
| | | | | | | | | | | | | | | | | | | | | |
DB 138 KRLRKFRNKIKKKIKG 155

RESULT 2
MP36.PIG STANDARD; PRT; 166 AA.
AC P49931;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Antibacterial peptide PMAP-36 precursor (Myeloid antibacterial peptide
DE 36).
GN PMAP36.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP TISSUE-BONE MARROW; PubMed=8293820;
RX MEDLINE=94123775; PubMed=8293820;
RA Storici P., Scocchi M., Tossi A., Gennaro R., Zanetti M.;
RT "Chemical synthesis and biological activity of a novel antibacterial
RT peptide deduced from a pig myeloid cDNA.";
RL FEBS Lett. 337:303-307(1994).
CC -1- FUNCTION: EXERTS ANTIMICROBIAL ACTIVITY AGAINST BOTH GRAM-POSITIVE
CC AND NEGATIVE BACTERIA. ITS ACTIVITY APPEARS TO BE MEDIATED BY ITS
CC ABILITY TO DAMAGE BACTERIAL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L29125; AAA31070.1; -.
CC InterPro: IPR001894; Cathelicidin.
CC Pfam: PF00666; Cathelicidins; 1.
CC ProDom: PD001838; Cathelicidins; 1.
CC PROSITE: PS00946; CATHELICIDINS_1; 1.
CC PROSITE: PS00947; CATHELICIDINS_2; 1.
CC Antibiotic; Signal.
CC FT SIGNAL 1 29 POTENTIAL.
CC FT PROPEP 30 129 ANTIMICROBIAL PEPTIDE PMAP-36.
CC FT CHAIN 130 166 PYRROLIDONE CARBOXYLIC ACID (BY
CC FT MOD_RES 30 30 SIMILARITY).
CC FT DISULFID 85 96 BY SIMILARITY.
CC FT DISULFID 107 124 BY SIMILARITY.
CC SEQUENCE 166 AA; 18647 MW; 94B13C69709DA64B CRC64;

Query Match 56.28; Score 50; DB 1; Length 166;
Best Local Similarity 52.9%; Pred. No. 4.7;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 RLKRFNRKIKKKIKG 18
| | | | | | | | | | | | | | | | | | | | | |
DB 132 RFRRLKRTKRLKKIG 148

RESULT 3
GRAM_MOUSE STANDARD; PRT; 173 AA.
ID GRAM_MOUSE
AC P51437;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cathelin-related antimicrobial peptide precursor (Cramp) (Cathelin-
DE like protein) (CLP).
GN CNLP OR CRAMP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Bone marrow;
RX MEDLINE=97294716; PubMed=9148921;
RA Gallo R.L., Kim K.D., Bernfield M., Kozak C.A., Zanetti M.,
RA Merluzzi L., Gennaro R.;
RT "Identification of CRAMP, a cathelin-related antimicrobial peptide
RT expressed in the embryonic and adult mouse.";
RL J. Biol. Chem. 272:13088-13093(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/C; TISSUE=Bone marrow;
RX MEDLINE=96326596; PubMed=8706928;
RA Popsueva A.E., Zinovjeva M.V., Vissers Y.W.M., Fibbe W.E.,
RA Belyavsky A.V.;
RT "A novel murine cathelin-like protein expressed in bone marrow.";
RL FEBS Lett. 391:5-8(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RA Hutterer K.M., Piraino J., Gallo R.L.;
RL submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ACTS AS A POTENT ANTIMICROBIAL PEPTIDE.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN TESTIS, SPLEEN, STOMACH, AND
CC INTESTINE. VERY LOW EXPRESSION FOUND IN HEART, LUNG AND SKELETAL
CC MUSCLE. NO EXPRESSION IN BRAIN, KIDNEY OR LIVER.
CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```


CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF FOUR PEPTIDE BONDS IN THE VIRAL
 CC PRECURSOR POLYPROTEIN, COMMONLY WITH ASP OR GLU IN THE P6
 CC POSITION. CYS OR THR IN P1 AND SER OR ALA IN P1'.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X03700; CAA27332.1; -
 DR PIR; A03914; GNMVY.
 DR HSSP; P14336; 1SVB.
 DR MEROPS; S07.001; -.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR000069; Flavi_M.
 DR InterPro; IPR001157; Flavi_NS1.
 DR InterPro; IPR000752; Flavi_NS2A.
 DR InterPro; IPR000487; Flavi_NS2B.
 DR InterPro; IPR000404; Flavi_NS4A.
 DR InterPro; IPR001528; Flavi_NS4B.
 DR InterPro; IPR000208; Flavi_NS5.
 DR InterPro; IPR001122; Flavi_capsid.
 DR InterPro; IPR000336; Flavi_glycoprote.
 DR InterPro; IPR001850; Flavi_helicase.
 DR InterPro; IPR002335; Flavi_propep.
 DR InterPro; IPR002877; FtsJ.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF01003; Flavi_capsid; 1.
 DR Pfam; PF00869; Flavi_glycoprot; 1.
 DR Pfam; PF02832; Flavi_glycop_C; 1.
 DR Pfam; PF00949; Flavi_helicase; 1.
 DR Pfam; PF01004; Flavi_M; 1.
 DR Pfam; PF00948; Flavi_NS1; 1.
 DR Pfam; PF01005; Flavi_NS2A; 1.
 DR Pfam; PF01002; Flavi_NS2B; 1.
 DR Pfam; PF01350; Flavi_NS4A; 1.
 DR Pfam; PF01349; Flavi_NS4B; 1.
 DR Pfam; PF00972; Flavi_NS5; 1.
 DR Pfam; PF01570; Flavi_propep; 1.
 DR Pfam; PF01728; FtsJ; 1.
 DR Pfam; PF00271; helicase_C; 1.
 DR ProDom; PD001496; Flavi_NS1; 1.
 DR ProDom; PD001556; Flavi_glycoprote; 1.
 DR SMART; SM00490; HELIC_C; 1.
 KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
 KW Core protein; Coat protein; Envelope protein; Hydrolase; Helicase;
 KW ATP-binding; Transmembrane; Nonstructural protein.
 FT INT_MET 1
 FT CHAIN 1 121
 FT PROPEP 122 210
 FT CHAIN 211 285
 FT CHAIN 286 778
 FT CHAIN 779 1130
 FT CHAIN 1131 1354
 FT CHAIN 1355 1484
 FT CHAIN 1485 2107
 FT CHAIN 2108 2256
 FT CHAIN 2257 2506
 FT CHAIN 2507 3411
 FT TRANSMEM 249 269
 FT TRANSMEM 271 285
 FT TRANSMEM 740 753
 FT TRANSMEM 755 778
 FT TRANSMEM 1159 1180
 FT DOMAIN 383 396

FT NF_BIND 1682 1689 ATP (POTENTIAL).
 FT SITE 1773 1776 DEAD BOX.
 FT DISULFID 288 315 BY SIMILARITY.
 FT DISULFID 345 401 BY SIMILARITY.
 FT DISULFID 359 390 BY SIMILARITY.
 FT DISULFID 377 406 BY SIMILARITY.
 FT DISULFID 467 568 BY SIMILARITY.
 FT DISULFID 585 615 BY SIMILARITY.
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 986 986 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2320 2320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2346 2346 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2467 2467 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 3411 AA; 379512 MW; 680E0FACD230CFA6 CRC64;
 Query Match 52.8%; Score 47; DB 1; Length 3411;
 Best Local Similarity 56.2%; Pred. No. 1.8e+02;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 3 LRKFRNKIKKLRKIG 18
 Db 20 VRSLSKIKKTKKQIG 35
 ID POLG_YEFPV2 STANDARD; PRT; 3411 AA.
 AC P19901;
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 40, Last annotation update)
 DE 16-OCT-2001 (Rel. 17, Last sequence update)
 DE Genome polyprotein (contains: Capsid protein C (Core protein); Matrix
 DE protein (Envelope protein M); Major envelope protein E; Nonstructural
 DE proteins NS1, NS2A, NS2B, NS4A and NS4B; Protease/helicase
 DE (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)
 DE (NS5)).
 OS yellow fever virus (strain Pasteur 17D-204).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Flavivirus.
 OX NCBI_TaxID=11091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89282413; Pubmed=2734112;
 RX Dupuy A., Despres P., Cahour A., Girard M., Bouloy M.;
 RT "Nucleotide sequence comparison of the genome of two 17D-204 yellow
 RT fever vaccines.";
 RL Nucleic Acids Res. 17:3989-3989(1989).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF FOUR PEPTIDE BONDS IN THE VIRAL
 CC PRECURSOR POLYPROTEIN, COMMONLY WITH ASP OR GLU IN THE P6
 CC POSITION. CYS OR THR IN P1 AND SER OR ALA IN P1'.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X15062; CAB37419.1; -
 DR PIR; S07757; GNMVY.
 DR HSSP; P14336; 1SVB.
 DR MEROPS; S07.001; -.

```
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000069; Flavi_M.
DR InterPro: IPR001157; Flavi_NSI.
DR InterPro: IPR000752; Flavi_NS2A.
DR InterPro: IPR000487; Flavi_NS2B.
DR InterPro: IPR000404; Flavi_NS4A.
DR InterPro: IPR001528; Flavi_NS4B.
DR InterPro: IPR000208; Flavi_NS5.
DR InterPro: IPR001122; Flavi_capsid.
DR InterPro: IPR000336; Flavi_glycoprote.
DR InterPro: IPR001850; Flavi_helicase.
DR InterPro: IPR002535; Flavi_propep.
DR InterPro: IPR002877; FtsJ.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF01003; Flavi_capsid.1.
DR Pfam: PF00869; Flavi_glycoprote.1.
DR Pfam: PF02832; Flavi_glycop_C.1.
DR Pfam: PF00949; Flavi_helicase.1.
DR Pfam: PF01004; Flavi_M.1.
DR Pfam: PF00948; Flavi_NSI.1.
DR Pfam: PF01005; Flavi_NS2A.1.
DR Pfam: PF01002; Flavi_NS2B.1.
DR Pfam: PF01350; Flavi_NS4A.1.
DR Pfam: PF01349; Flavi_NS4B.1.
DR Pfam: PF00972; Flavi_NS5.1.
DR Pfam: PF01570; Flavi_propep.1.
DR Pfam: PF01728; FtsJ.1.
DR Pfam: PF00271; helicase_C.1.
DR ProDom: PD001496; Flavi_NSI.1.
DR SMART: SM00490; HELIC_C.1.
DR PolyProtein: Glycoprotein; Transferrase; RNA-directed RNA polymerase;
KM Core protein; Coat protein; Envelope protein; Hydrolyase; Helicase;
KM ATP-binding; Transmembrane; Nonstructural protein.
FT INIT_MET 1
FT CHAIN 1 121
FT PROPEP 122 210
FT CHAIN 211 285
FT CHAIN 286 778
FT CHAIN 779 1130
FT CHAIN 1131 1354
FT CHAIN 1355 1484
FT CHAIN 1485 2107
FT CHAIN 2108 2256
FT CHAIN 2257 2506
FT CHAIN 2507 3411
FT NP_BIND 1682 1689
FT SITE 1773 1776
FT TRANSMEM 249 269
FT TRANSMEM 271 285
FT TRANSMEM 740 753
FT TRANSMEM 755 778
FT TRANSMEM 1159 1180
FT DISULFID 288 315
FT DISULFID 345 401
FT DISULFID 359 390
FT DISULFID 377 406
FT DISULFID 467 568
FT DISULFID 585 615
FT CARBOHD 1134 134
FT CARBOHD 130 150
FT CARBOHD 908 908
FT CARBOHD 986 986
FT CARBOHD 2330 2330
FT CARBOHD 2346 2346
FT CARBOHD 2467 2467
SQ SEQUENCE 3411 AA: 379524 MM: 3298C0771FED23F7 CRC64;
```

Query Match 52.8%; Score 47; DB 1; Length 3411;
Best Local Similarity 56.2%; Pred. No. 1.8e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```
QY 3 LRKFRNKIKKKKIG 18
Db 20 VRSLSNKKIKOKTKOIG 35

RESULT 9
DEF_VIBCH STANDARD; PRT; 169 AA.
ID DEF_VIBCH
AC G9KV03;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR VC0046.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxId=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EU TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emdolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterlind T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
CC -!- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (by similarity).
CC -!- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -!- COFACTOR: Binds 1 iron(II) ion (by similarity).
CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
```

Query Match 51.7%; Score 46; DB 1; Length 169;
Best Local Similarity 81.8%; Pred. No. 16;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 RNRKIKERLKI 17
Db 152 RNRKIKERLKI 162
RESULT 10

```

SECA_AQUAE
ID SECA_AQUAE STANDARD: PRT; 984 AA.
AC 067718;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Preprotein translocase SECA subunit.
GN SECA OR AQ_1870.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Decerit G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjaj M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson M.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus."
RL Nature 392:353-358(1998).
CC -I- FUNCTION: INVOLVED IN PROTEIN EXPORT. INTERACTS WITH THE SECA/SECE
CC SUBUNITS. SECA HAS A CENTRAL ROLE IN COUPLING THE HYDROLYSIS OF
CC ATP TO THE TRANSFER OF PRE-SECRETORY PERIPLASMIC AND OUTER
CC MEMBRANE PROTEINS ACROSS THE MEMBRANE (BY SIMILARITY).
CC -I- SUBUNIT: PART OF THE PROKARYOTIC PROTEIN TRANSLOCATION APPARATUS
CC WHICH COMPRISE SECA, SECB, SECD, SECE, SECF, SECG AND SECY
CC (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF PLASMA MEMBRANE
CC (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE SECA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE000760; AAC07677.1; -.
DR InterPro: IPR000185; SECA.
DR Pfam: PF01043; SECA_Protein; 2.
DR PRINTS: PR00906; SECA.
DR PROSITE: PS01312; SECA; 1.
DR Protein transport; ATP-binding; Membrane; Translocation; Transport;
DR Complete proteome.
DR NP_BIND 111
DR SEQUENCE 984 AA; 113976 MW; AE23834D30DED84 CRC64;
SO
Query Match 51.7%; Score 46; DB 1; Length 984;
Best Local Similarity 58.8%; Pred. No. 79;
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 KRLRKFRNKIKKELKRI 17
Db 19 KRLRKFRNKIKKELKRI 35
RESULT 11
PPOL_SARPE STANDARD: PRT; 996 AA.
AC 011208;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Poly (ADP-ribose) polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)-ADP-
DE ribosyltransferase) (POLY(ADP-ribose) synthetase).
OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Sarcophagidae; Sarcophaga.

```

```

OX NCBI_TaxID=7386;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=94170813; PubMed=8125121;
RA Masutani M., Nozaki T., Hltoni Y., Ikejima M., Nagasaki K.,
RA de Prati A.C., Kurata S., Natori S., Sugimura T., Esumi H.;
RT "Cloning and functional expression of poly(ADP-ribose) polymerase
RT cDNA from Sarcophaga peregrina."
RL Eur. J. Biochem. 220:607-614(1994).
CC -I- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC -I- CATALYTIC ACTIVITY: NAD(+) + {ADP-ribosyl}(n)-acceptor =
CC nicotinamide + {ADP-D-ribosyl}(n+1)-acceptor.
CC -I- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC -I- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC -I- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D16482; BAA0943.1; -.
DR HSSP: P26446; 1A26.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR InterPro: IPR004102; PARP_reg.
DR InterPro: IPR001510; Znf-PARP.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR Pfam: PF02877; PARP_reg; 1.
DR Pfam: PF00645; zfp-PARP; 2.
DR ProDom: PD004675; Znf-PARP; 2.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS50172; BRCT; 1.
DR PROSITE: PS00347; PARP_ZN_FINGER_1; FALSE_NEG.
DR PROSITE: PS50064; PARP_ZN_FINGER_2; 2.
DR Transferase; Glycosyltransferase; NAD: DNA-binding; Nuclear protein;
KW ADP-ribosylation; Zinc-finger; Zinc.
FT DNA_BIND 1 369 BY SIMILARITY.
FT DOMAIN 370 507 AUTOMODIFICATION DOMAIN.
FT DOMAIN 382 456 BRCT.
FT DOMAIN 508 996 NAD-BINDING.
FT ZN_FING 126 164 BY SIMILARITY.
FT ZN_FING 211 214 NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT DOMAIN 232 235 NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT SEQUENCE 996 AA; 113018 MW; 690DD36E7487298 CRC64;
SO
Query Match 51.7%; Score 46; DB 1; Length 996;
Best Local Similarity 56.2%; Pred. No. 80;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 KRLRKFRNKIKKELKRI 16
Db 238 KRLRKFRNKIKKELKRI 253
RESULT 12

```

YJEP	HA2IN	STANDARD:	PRT: 1111 AA.
ID	YJEP_HA2IN		
AC	O57362.1		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Protein HI0195.1 precursor.		
GN	HI0195.1.		
OS	Haemophilus influenzae.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		
OC	Haemophilus.		
OX	NCBI_taxid=727;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=RD / KM20 / ATCC 51907.		
RX	MEDLINE=95350630. PubMed=7542800:		
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,		
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,		
RA	McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,		
RA	Scott J.D., Shirley R., Liu L.-I., Godke A., Kelley J.M.,		
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,		
RA	Uterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C.,		
RA	Fine L.D., Fitcham J.A., Fuhrman J.L., Geoghagen N.S.M.,		
RA	Gnam C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,		
RA	Venter J.C.;		
RT	"Whole-genome random sequencing and assembly of Haemophilus		
RT	influenzae Rd.";		
RL	Science 269:496-512(1995).		
RN	[2]		
RP	IDENTIFICATION BY MASS SPECTROMETRY.		
RX	MEDLINE=20137488. PubMed=10675023:		
RA	Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wiipf B.,		
RA	Gray C., Pountoulakis M.;		
RT	"Two-dimensional map of the proteome of Haemophilus influenzae.";		
RL	Electrophoresis 21:411-429(2000).		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).		
CC	-1- SIMILARITY: BELONGS TO THE UPF0003 FAMILY. STRONG, TO E.COLI YJEP.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on ways		
CC	the European Bioinformatics Institute. There are no restrictions on ways		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: U32704. AAC21864.1; -		
DR	TIGR: HI0195.1; -		
DR	InterPro: IPR001880; UPF00003.		
DR	Pfam: PF00924; MS_channel; 1.		
DR	PROSITE: PS01246; UPF0003; 1.		
KW	transmembrane; Signal; Complete proteome.		
FT	SIGNAL	1	31
FT	CHAIN	32	1111
FT	TRANSMEM	490	510
FT	TRANSMEM	538	558
FT	TRANSMEM	572	592
FT	TRANSMEM	620	640
FT	TRANSMEM	644	664
FT	TRANSMEM	694	714
FT	TRANSMEM	797	817
FT	TRANSMEM	840	860
FT	TRANSMEM	885	905
FT	TRANSMEM	922	942
FT	TRANSMEM	1003	1023
FT	SEQUENCE	1111 AA;	126825 MW; 15DF8F9F7E4E701C CRC64;

Query Match	51.7%	Score 46;	DB 1;	Length 111;
Best Local Similarity	61.5%	Pred. No. 89;		
Matches	8;	Conservative	4;	Mismatches 1; Indels 0; Gaps 0;
0y	5	KFRNKIKKKLKI	17	

```

Db          509      KFKNRIRKQOLNKI 521

| | : | | : | |
| | : | | : | |

RESULT 13
MYG_PHOSI
ID MYG_PHOSI STANDARD; PRT; 153 AA.
AC P30562;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myoglobin.
GN MB.
OS Phoca sibirica (Baikal seal).
OC Eukaryota; Metazoa; Chordata; Cranialta; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Plinnipedia; Prociidae; Phoca.
OX NCBI_TaxId=9719;
[ ]
SEQUENCE
MEDLINE=92222476;. Pubmed=1807253;
RA Batam G.I., Grachev M.A., Malikov N.G., Nazimov I.V., Shemyakin V.
RT "Amino acid sequence of myoglobin from seals from Lake Baikal." ;
RL Bloorq. Khim. 17:1166-1171(1991).
CC -1- FUNCTION. SERVES AS A RESERVE SUPPLY OF OXYGEN AND FACILITATES
CC THE MOVEMENT OF OXYGEN WITHIN MUSCLES.
CC CC
CC -1- SIMILARITY. BELONGS TO THE GLOBIN FAMILY.
CC PIR: JN0344; JN0344.
DR HSSP; P02162; IMBS.
DR InterPro; IPR000971; Globin.
DR InterPro; IPR002335; Myoglobin.
DR Pfam; PF00042; globin. 1.
DR PRINTS; PR00613; MYOGLOBIN.
DR PROSITE; PS01033; GLOBIN; 1.
KT Heme; Oxygen transport; Transport; Muscle.
FM METAL 64 64 IRON (HEME DISTAL LIGAND).
FT METAL 93 93 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 153 AA; 17952 MW; E126DAV57BDEBD9 CRC64;
```

```

Query Match 50.6%; Score 45; DB 1; Length 153;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY      1 KRLKFRNKKIKELKKIG 18
          | | | | | | | |
DB      133 KALELFRNDIAKIKELG 150

RESULT 14
SYS_AERPE
ID     SYS_AERPE      STANDARD;      PRT;      463 AA.
AC     O9YAG3:
DT     16-OCT-2001 (Rel. 40, Created)
DT     16-OCT-2001 (Rel. 40, Last sequence update)
DT     16-OCT-2001 (Rel. 40, Last annotation update)
DE     Seryl-tRNA synthetase (EC 6.1.1.11) (Seryl--tRNA ligase) (SERS).
SERS   ON APEL1976.
OS     Aeropyrum pernix.
OC     Archaeae: Crenarchaeota: Desulfurococcales; Desulfurococcaceae;
OC     Aeropyrum.
OX     NCBI_TaxID=56636;
RN     (1)
RP     SEQUENCE FROM N.A.
RC     STRAIN-K1;
RX     MEDLINE=99310339; PubMed=10382966;
RA     Kavarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA     Jin-no K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,
RA     Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA     Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA     Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA     Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT     "Complete genome sequence of an aerobic hyper-thermophilic
RT     crenarchaeon, Aeropyrum pernix K1."

```

RL DNA Res. 6:83-101(1999).
 CC -I- CATALYTIC ACTIVITY: ATP + L-serine + tRNA(Ser) = AMP + diphosphate
 CC + L-seryl-tRNA(Ser).
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -I- SIMILARITY: BELONGS TO CLASS-III AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP000063; BAA0986.1; -
 DR HSSP: P34945.1SER.
 DR InterPro: IPR002106; AA_tRNA_ligase_II.
 DR InterPro: IPR002314; tRNA-synt_2b.
 DR InterPro: IPR002317; tRNA-synt_ser.
 DR Pfam: PF00587; tRNA-synt_2b; 1.
 DR PRINTS: PR00981; TRNASYNTHSER.
 DR PROSITE: PS00179; AA_tRNA_LIGASE_II_1; 1.
 DR PROSITE: PS00339; AA_tRNA_LIGASE_II_2; FALSE_NEG.
 DR K06H7.3.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 463 AA; 53596 MW; 12126FE871A22FB CRC64;

Query Match 50.6%; Score 45; DB 1; Length 463;
 Best Local Similarity 52.9%; Pred. No. 54;
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 KRLKFRNKIKKIKKKI 17
 |||:|:|||||:
 Db 82 KRLKREDLEKIKKKV 98

RESULT 15
 YMX3_CAEEL
 ID YMX3_CAEEL STANDARD: PRT; 618 AA.
 AC P34511;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 71.3 kDa protein K06H7.3 in chromosome III.
 GN K06H7.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen M.,
 RA Sims M., Smaison N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 368:32-38(1994).
 RN [2]
 RP REVISIONS.
 RA Waterston R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -I- SIMILARITY: TO ISOPEPTENYL-DIPHOSPHATE DELTA ISOMERASE.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: LJ5314; AAF9988.2; -
 DR PIR: S44843; S44843.
 DR WormPep: K06H7.3; CE25044.
 DR InterPro: IPR000086; NUDIX_hydrolase.
 DR InterPro: IPR000822; znf-C2H2.
 DR Pfam: PF00293; NUDIX; 1.
 DR ProDom: PD004109; Isopeptidiph_delta_isomerase; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 618 AA; 71320 MW; DE6A695E14F6EAE4 CRC64;

Query Match 50.6%; Score 45; DB 1; Length 618;
 Best Local Similarity 62.5%; Pred. No. 70;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KRLKFRNKIKKIKKK 16
 |:|:|:|||||:
 Db 527 KKEKKARQKEKEKKKK 542

Search completed: July 12, 2002, 08:04:42
 Job time: 654 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:03:51 ; Search time 174.7 Seconds
(without alignments)
17.824 Million cell updates/sec

Title: US-09-642-744B-19
Perfect score: 89
Sequence: 1 KRLKFRNKIKERLKTIG 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_prodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	56.2	178	4	060831
2	50	56.2	178	11	09JIG8
3	49	55.1	142	16	066881
4	49	55.1	324	10	042621
5	49	55.1	326	10	042622
6	48	53.9	381	17	0978N0
7	48	53.9	439	2	053589
8	48	53.9	497	16	099873
9	48	53.9	598	2	048537
10	47.5	53.4	274	17	028716
11	47	52.8	121	16	067804
12	47	52.8	141	5	09NH7
13	47	52.8	257	12	091E33
14	47	52.8	400	12	089285
15	47	52.8	400	12	089297
16	47	52.8	400	12	089299

17	47	52.8	400	12	089304	089304 yellow feve
18	47	52.8	400	12	089311	089311 yellow feve
19	47	52.8	400	12	089315	089315 yellow feve
20	47	52.8	400	12	089317	089317 yellow feve
21	47	52.8	778	12	09W9B8	09W9B8 yellow feve
22	47	52.8	778	12	P89915	P89915 yellow feve
23	47	52.8	3411	12	098803	098803 yellow feve
24	47	52.8	3411	12	091857	091857 yellow feve
25	47	52.8	3411	12	09YWN2	09YWN2 yellow feve
26	47	52.8	3411	12	09YWN1	09YWN1 yellow feve
27	47	52.8	3411	12	09YWN0	09YWN0 yellow feve
28	47	52.8	3411	12	09YRV3	09YRV3 yellow feve
29	47	52.8	3411	12	089275	089275 yellow feve
30	47	52.8	3411	12	089277	089277 yellow feve
31	47	52.8	3411	12	089278	089278 yellow feve
32	47	52.8	3411	12	089276	089276 yellow feve
33	46	51.7	188	4	075915	075915 homo sapien
34	46	51.7	188	11	09DB37	09DB37 mus musculu
35	46	51.7	188	11	09ES40	09ES40 ratu
36	46	51.7	259	10	09M5W3	09M5W3 euphorbia e
37	46	51.7	305	5	045706	045706 caenorhabdi
38	46	51.7	329	10	042620	042620 brassica na
39	46	51.7	331	10	042619	042619 brassica na
40	46	51.7	400	12	089308	089308 yellow feve
41	46	51.7	400	12	089309	089309 yellow feve
42	46	51.7	444	5	022405	022405 caenorhabdi
43	46	51.7	1013	3	09C130	09C130 picchia etch
44	45.5	51.1	161	8	095BN8	095BN8 hemitomes c
45	45.5	51.1	637	10	09L100	09L100 arabidopsis

ALIGNMENTS

RESULT 1

ID 060831 PRELIMINARY; PRT: 178 AA.

AC 060831: 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE JM4 PROTEIN.

GN JM4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Strom T.M., Nyakatura G., Hellebrand H., Drescher B., Rosenthal A., Meindl A.,

RT "Transcription map in Xp11.23."

RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ005896; CA06753.1; -

SQ SEQUENCE 178 AA; 19258 MW; E62695D7DC40550A CRC64;

Query Match 56.2%; Score 50; DB 4; Length 178;
Best Local Similarity 52.9%; Pred. No. 20;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 KRLKFRNKIKERLKTIG 18
DB 141 KRLKFRNKIKERLKTIG 157

RESULT 2

ID 09JIG8 PRELIMINARY; PRT: 178 AA.

AC 09JIG8: 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE DXIMX39E PROTEIN (DNA SEGMENT, CHR X, IMGNEX 39, EXPRESSED).
GN DXIMX39E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20313888; PubMed=10857745;
RA Means G.D., Toy D.Y., Baum P.R., Derry J.M.J.;
RT "A transcript map of a 2-Mb BAC contig in the proximal portion of the
mouse X chromosome and regional mapping of the scurfy mutation.";
RL Genomics 65:213-223(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,
RA Brownstein M.D., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyszewski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
DR EMBL: AF229636; AAF66950.1; -;
DR EMBL: AK003344; BAB22729.1; -;
DR MGD: MGI:1859607; DXIMX39E.
SQ SEQUENCE 178 AA: 19478 MW: 48DE686D480D70DD CRC64;

Query Match 56.2%; Score 50; DB 11; Length 178;
Best Local Similarity 52.9%; Pred. No. 20;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 RLKFRNKIKELKTKIG 18
DB 141 RLRLNKKIKENKTESIG 157

RESULT 3
066881 PRELIMINARY: PRT; 142 AA.
ID 066881;
AC 066881;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL 16.4 KDA PROTEIN.
GN AQ.635.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";

RL Nature 392:353-358(1998).
DR EMBL: AE000699; AAC06842.1; -;
DR InterPro: IPR000106; Low_mwt_PTase.
DR Pfam: PF01451; LMPec; 1.
DR SMART: SM00226; LMPec; 1.
KW Hypothetical protein; complete proteome.
SQ SEQUENCE 142 AA: 16403 MW: C3AF0A2783FF145 CRC64;

Query Match 55.1%; Score 49; DB 16; Length 142;
Best Local Similarity 60.0%; Pred. No. 22;
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 3 RLKFRNKIKELKTKI 17
DB 115 LRLRLKIEELKTKL 129

RESULT 4
042621 PRELIMINARY: PRT; 324 AA.
ID 042621;
AC 042621;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CTP:PHOSHOCHOLINE CYTIDYLTRANSFERASE (EC 2.7.7.15).
GN CCT.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Brassica.
NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JET NURF; TISSUE=ROOT;

Query Match 55.1%; Score 49; DB 10; Length 324;
Best Local Similarity 47.1%; Pred. No. 48;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLKFRNKIKELKTKIG 18
DB 196 RLKLOEKVKEOEKVG 212

RESULT 5
042622 PRELIMINARY: PRT; 326 AA.
ID 042622;
AC 042622;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CTP:PHOSHOCHOLINE CYTIDYLTRANSFERASE (EC 2.7.7.15).
GN CCT.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Brassica.
NCBI_TaxID=3708;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JET NURF; TISSUE=ROOT;

RA Nishida T., Swincoe R., Slabas A.R., Murata N.,
RT "Cloning of *Brassica napus* CTP:phosphocholine cytidyltransferase
RT cDNAs by complementation of the in a yeast mutant.",
RL Plant Mol. Biol. 0:0-0(1996).
DR EMBL: D63168; BAA09644.1; .
DR HSSP: P27623; 1COZ.
DR InterPro: IPR001994; Cytidyltransferase.
DR Pfam: PF01467; Cytidyltransferase; 1.
DR Trnaseq: Nucleotidyltransferase.
KW TRANSFERASE
SQ SEQUENCE 326 AA; 38024 MW; 405C7E054C18196F CRC64;

Query Match 55.1%; Score 49; DB 10; Length 326;
Best Local Similarity 47.1%; Pred. No. 48;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 2 RLKFRNKIKKIKKIG 18
DB 198 RLKLUQKYEEOEKVG 214

RESULT 6
O978N0 PRELIMINARY; PRT; 381 AA.
AC O978N0;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE IAA-AMINO ACID HYDROLASE.
GN TVG4333041.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
OC Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GSSI / DSM 4299 / JCM 9571;
RA MEDLINE=20570465; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima-Ohya Y., Watanabe K., Yamazaki K., Kanehori K., Kawamoto T.,
RA Nunoshima T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of *Thermoplasma volcanium*,"
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
DR EMBL: AP00096; BAB60527.1; .
DR InterPro: IPR002933; Peptidase_M20.
DR Pfam: PF01546; Peptidase_M20; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 381 AA; 42168 MW; EF1956GB2C0FA949 CRC64;

Query Match 53.9%; Score 48; DB 17; Length 381;
Best Local Similarity 69.2%; Pred. No. 75;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 KRFNRKIKKIKK 16
DB 369 KRFNRKIKKIKK 381

RESULT 7
O53589 PRELIMINARY; PRT; 439 AA.
AC O53589;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORFS 1,2,3 & 4.
OS *Staphylococcus aureus*.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-NEWMAN;
RA Wieland K.P., Goetz F.;
RT "Regulation of the staphyloxanthin biosynthesis in *Staphylococcus*
RT *aureus* Newman.",
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: X97985; CAA66626.1; .
DR InterPro: IPR001171; Bac_phytoene_dh.
DR Pfam: PF02032; Phytoene_dh; 1.
KW PHYTOENE
SQ SEQUENCE 439 AA; 50829 MW; AB03097B55E83862 CRC64;

Query Match 53.9%; Score 48; DB 2; Length 439;
Best Local Similarity 64.3%; Pred. No. 86;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 5 KFRDKIKKIKKIG 18
DB 393 KFRDKIKKIKKIG 406

RESULT 8
O99R73 PRELIMINARY; PRT; 497 AA.
AC O99R73;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SA2351 PROTEIN (HYPOTHETICAL PROTEIN SAV2564).
GN SA2351 OR SAV2564.
OS *Staphylococcus aureus* (strain N315), and
OS *Staphylococcus aureus* (strain Mu50).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=158879, 158878;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S.aureus (strain N315), and S.aureus (strain Mu50);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-U I., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant *Staphylococcus*
RT *aureus*,"
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003137; BAB43655.1; .
DR EMBL: AP003137; BAB58726.1; .
DR InterPro: IPR000171; Bac_phytoene_dh.
DR Pfam: PF02032; Phytoene_dh; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 497 AA; 57174 MW; E20BB9DDF514C9D CRC64;

Query Match 53.9%; Score 48; DB 16; Length 497;
Best Local Similarity 64.3%; Pred. No. 96;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 5 KFRNKIKKIKKIG 18
DB 393 KFRDKIKKIKKIG 406

RESULT 9
O48537 PRELIMINARY; PRT; 598 AA.
AC O48537;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
RN [1]

```
DE DNA PRIMASE.
GN ORE1.
OS Lactobacillus delbrueckii.
OC Plasmid pMS58.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1584;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS58;
RA Klein J.R., Heinrich B.;
RT "Complete sequence of the cryptic plasmid pMS58 from Lactobacillus
RL delbrueckii subsp. lactis."
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z50864; CAA90741.1; -.
KW Plasmid.
SQ SEQUENCE 598 AA; 68783 MW; FF65175251ED76F2 CRC64;

Query Match 53.9%; Score 48; DB 2; Length 598;
Best Local Similarity 41.2%; Pred. No. 1.le+02;
Matches 7; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 2 LRKRFNRKIKKIKIG 18
| : : : : : : : : : :
DB 529 RPRFRKMKIKOSVKKIG 545

RESULT 10
028716 PRELIMINARY; PRT; 274 AA.
AC 028716;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-AUG-1998 (TREMBLrel. 05, Last sequence update)
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN AF1556.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervavage A.R., Graham D.E., Kyprides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
DR EMBL; AE000995; AAB89691.1; -.
DR TIGR; AF1556; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 274 AA; 30955 MW; F2FBD4F9D1F2BF13 CRC64;
```

```
RESULT 11
067804 PRELIMINARY; PRT; 121 AA.
AC 067804;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 14.6 KDA PROTEIN.
GN AQ_2000.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
DR EMBL; AE000767; AAC07774.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 121 AA; 14642 MW; C5C6FA762E0CD0C6 CRC64;
```

```
Query Match 52.8%; Score 47; DB 16; Length 121;
Best Local Similarity 58.8%; Pred. No. 36;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 KRLKRFNRKIKKIKKI 17
| : : : : : : : : : :
DB 28 KNLSEFRKIKENSKKI 44
```

```
RESULT 12
09NNH7 PRELIMINARY; PRT; 141 AA.
AC 09NNH7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PRFMP1 PROTEIN (FRAGMENT).
GN V4R.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3;
RX MEDLINE=20520970; PubMed=11071291;
RA Taylor H.M., Kyes S.A., Newbold C.I.;
RT "Var gene diversity in Plasmodium falciparum is generated by frequent
RT recombination events."
RL Mol. Biochem. Parasitol. 110:391-397(2000).
DR EMBL; AF221813; AAF3653.1; -.
FT NON_TER 1
FT NON_TER 141
SQ SEQUENCE 141 AA; 16868 MW; 7987FE800C954BA CRC64;
```

```
Query Match 52.8%; Score 47; DB 5; Length 141;
Best Local Similarity 58.8%; Pred. No. 41;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 KRLKRFNRKIKKIKKI 17
| : : : : : : : : : :
DB 21 KRLKRFNRKIKKIKKI 37
```

```
RESULT 13
```

```
091E33
ID 091E33 PRELIMINARY; PRT; 257 AA.
AC 091E33;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE 28 kDa PROTEIN.
OS Pepper mild mottle virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxId=12239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TA;
RA Velasco L., Janssen D., Ruiz-Garcia L., Segundo E., Cuadrado I.M.;
RT "The complete nucleotide sequence of the RNA genome of a Pepper Mild
RT Mottle Virus (PMOV) isolate that overcomes L3 resistance in pepper.";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ308228; CAC5957.1; -.
SQ SEQUENCE 257 AA; 28514 MW; 66EDE458B04DB3B7 CRC64;

Query Match 52.8%; Score 47; DB 12; Length 257;
Best Local Similarity 60.0%; Pred. NO. 71;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LRKFRNKIKKKIG 16
DB 206 LRKFRNKQAKKRSNK 220

RESULT 14
ID 089295 PRELIMINARY; PRT; 400 AA.
AC 089295;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Yellow fever virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxId=11089;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRACIL35;
RX MEDLINE=97078788; PubMed=8918913;
RA Wang E., Weaver S.C., Shope R.E., Tesh R.B., Watts D.M., Barrett A.D.;
RT "Genetic variation in yellow fever virus: duplication in the 3'
RT noncoding region of strains from Africa.";
RL Virology 225:274-281(1996).
DR EMBL: U52389; AAB01953.1; -.
DR InterPro: IPR001122; Flavi_capsid.
DR InterPro: IPR000336; Flavi_glycoprote.
DR InterPro: IPR000069; Flavi_M.
DR Pfam: PF01003; Flavi_propep.
DR Pfam: PF00869; Flavi_glycoprot.
DR Pfam: PF01004; Flavi_M.
DR Pfam: PF01570; Flavi_propep.
DR ProDom: PD001556; Flavi_glycoprote.
FT NON_TER 400
SQ SEQUENCE 400 AA; 44278 MW; ADAC0A8E44EF877C CRC64;

Query Match 52.8%; Score 47; DB 12; Length 400;
Best Local Similarity 56.2%; Pred. NO. 11e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
RESULT 15
ID 089297 PRELIMINARY; PRT; 400 AA.
AC 089297;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Yellow fever virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxId=11089;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CAR77(83);
RX MEDLINE=97078788; PubMed=8918913;
RA Wang E., Weaver S.C., Shope R.E., Tesh R.B., Watts D.M., Barrett A.D.;
RT "Genetic variation in yellow fever virus: duplication in the 3'
RT noncoding region of strains from Africa.";
RL Virology 225:274-281(1996).
DR EMBL: U52392; AAB01955.1; -.
DR InterPro: IPR001122; Flavi_capsid.
DR InterPro: IPR000336; Flavi_glycoprote.
DR InterPro: IPR000069; Flavi_M.
DR Pfam: PF01003; Flavi_propep.
DR Pfam: PF00869; Flavi_glycoprot.
DR Pfam: PF01004; Flavi_M.
DR Pfam: PF01570; Flavi_propep.
DR ProDom: PD001556; Flavi_glycoprote.
FT NON_TER 400
SQ SEQUENCE 400 AA; 44491 MW; 69191A40EF674948 CRC64;

Query Match 52.8%; Score 47; DB 12; Length 400;
Best Local Similarity 56.2%; Pred. NO. 11e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 3 LRKFRNKIKKKIG 18
DB 20 VRSLSNKKIKKTKOIG 35
```

Search completed: July 12, 2002, 08:03:53
Job time: 639 sec

```
Query Match 52.8%; Score 47; DB 12; Length 400;
Best Local Similarity 56.2%; Pred. NO. 11e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LRKFRNKIKKKIG 18
DB 20 VRSLSNKKIKKTKOIG 35
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:00:44 ; Search time 227.32 Seconds
(without alignments)
9.284 Million cell updates/sec

Title: US-09-642-744B-21

Perfect score: 93

Sequence: 1 LRKFRNKIKKELKKIGQKI 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : Maximum Match 100%

Listing first 45 summaries

```

1: A_Geneseq_032802.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	19	22	AAB70669
2	93	100.0	21	22	AAB70670
3	93	100.0	22	22	AAB70668
4	93	100.0	32	21	AAB28486
5	93	100.0	37	22	AAB70665
6	93	100.0	38	22	AAB51193
7	87	93.5	29	15	AA45671
8	87	93.5	29	21	AAB07903
9	79	84.9	18	22	AAB70667
10	79	84.9	21	22	AAB70666
11	75	80.6	15	22	AAB70671

12	64	68.8	33	17	AA94449	Mouse antimicrobia
13	62	66.7	15	22	AAB70672	Lupine RCAP 18 cat
14	62	66.7	17	22	AAB70673	Lupine RCAP 18 cat
15	58	62.4	36	22	AA666421	Mouse cathelin pep
16	55	59.1	16	15	AA45673	RNP heparin bindi
17	55	59.1	16	21	AAB07905	Heparin-binding se
18	49	52.7	247	21	AA606758	Arabidopsis thalia
19	49	52.7	247	21	AA632601	Arabidopsis thalia
20	49	52.7	332	21	AA606757	Arabidopsis thalia
21	49	52.7	332	21	AA632600	Arabidopsis thalia
22	49	52.7	3421	22	AA884902	Chimeric protein f
23	48	51.6	178	22	AA894612	Human protein sequ
24	48	51.6	439	22	AAU34200	Staphylococcus aur
25	47	50.5	18	20	AA57472	Antimicrobial pept
26	47	50.5	157	21	AA629943	Arabidopsis thalia
27	47	50.5	174	21	AA629942	Arabidopsis thalia
28	47	50.5	183	21	AA629941	Arabidopsis thalia
29	47	50.5	189	21	AA624928	Syntaxin 1A N-term
30	47	50.5	216	18	AA624932	Syntaxin 1A delet
31	47	50.5	236	18	AA624933	Syntaxin 1A delet
32	47	50.5	264	18	AA624931	Syntaxin 1A N-term
33	47	50.5	266	18	AA624930	Syntaxin 1A N-term
34	47	50.5	285	19	AA643418	Rat syntaxin 1A C-
35	47	50.5	288	17	AA696421	Rat syntaxin 1A.
36	47	50.5	288	18	AA630105	Rat syntaxin 1A.
37	47	50.5	288	18	AA624927	Rat syntaxin 1A.
38	47	50.5	288	21	AA612822	Rat syntaxin 1A am
39	47	50.5	288	22	AA600254	SNARE homologue, h
40	47	50.5	857	22	AA663922	Drosophila melanog
41	46	49.5	27	16	AA684142	Peptide enhancer o
42	46	49.5	27	16	AA677056	Synthetic anti-neo
43	46	49.5	27	16	AA674705	Tryptic digestion
44	46	49.5	27	16	AA664784	Amphipathic peptid
45	46	49.5	27	17	AA692386	Lytic peptide used

ALIGNMENTS

RESULT 1
AAB70669 standard; peptide; 19 AA.

XX AAB70669:

XX 15-MAY-2001 (first entry)

XX Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:22.

XX Ovine; SMAR29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.

XX Oryctolagus cuniculus.

XX WO200112668-A1.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-US22781.

XX 18-AUG-1999; 9905-0149886.

XX (IOWA) UNIV IOWA RES FOUND.

XX (REGC) UNIV CALIFORNIA.

XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R.

XX WPI; 2001-234911/24.

XX New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections .

XX	Claim 1; Page 103; 137pp; English.
XX	
CC	AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC	AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC	SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC	derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC	antibiotic, antimicrobial and antiviral activities, and can be used as
CC	microbial growth and proliferation inhibitors and in gene therapy. (I)
CC	are useful for inhibiting microbial growth in an environment capable of
CC	sustaining such growth, for inhibiting microbial growth or strain in a
CC	host, and inhibiting the growth of drug-resistant microbial strains such
CC	as <i>Pseudomonas aeruginosa</i> , <i>Burkholderia cepacia</i> , <i>Alcaligenes</i> and
CC	<i>Xanthomonas</i> .
CC	
XX	
SO	Sequence 19 AA;
XX	
XX	
XX	Query Match 100.0%; Score 93; DB 22; Length 19;
XX	Best Local Similarity 100.0%; Pred. No. 1,1e-06;
XX	Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 LKFKNKIKKKIKKIGQKI 19
Db	1 LKFKNKIKKKIKKIGQKI 19
XX	
XX	RESULT 2
ID	AAB70670
XX	ID AAB70670 standard; peptide; 21 AA.
XX	
AC	AAB70670;
XX	
DT	15-MAY-2001 (first entry)
XX	
DE	Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:23.
XX	
XX	
KW	Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW	bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW	proliferation inhibitor; gene therapy; <i>Pseudomonas aeruginosa</i> ;
KW	<i>Burkholderia cepacia</i> ; <i>Alcaligenes</i> ; <i>Xanthomonas</i> .
XX	
OS	<i>Oryctolagus cuniculus</i> .
XX	
PN	MO200112668-A1.
XX	
PD	22-FEB-2001.
XX	
PF	18-AUG-2000; 2000MO-US22781.
XX	
PR	18-AUG-1999; 99US-0149886.
XX	
PA	(IOWA) UNIV IOWA RES FOUND.
PA	(RESC) UNIV CALIFORNIA.
XX	
PI	Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX	
DR	WPI: 2001-234911/24.
XX	
PT	New antimicrobial peptides useful as antibiotics for inhibiting growth
PT	and proliferation of microbes, and for treating microbial infections -
XX	
PS	Claim 1; Page 103; 137pp; English.
XX	
CC	AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC	AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC	SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC	derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC	antibiotic, antimicrobial and antiviral activities, and can be used as
CC	microbial growth and proliferation inhibitors and in gene therapy. (I)
CC	are useful for inhibiting microbial growth in an environment capable of
CC	sustaining such growth, for inhibiting microbial growth or strain in a
CC	host, and inhibiting the growth of drug-resistant microbial strains such
CC	as <i>Pseudomonas aeruginosa</i> , <i>Burkholderia cepacia</i> , <i>Alcaligenes</i> and
CC	<i>Xanthomonas</i> .
CC	
XX	
SO	Sequence 19 AA;
XX	
XX	
XX	Query Match 100.0%; Score 93; DB 22; Length 19;
XX	Best Local Similarity 100.0%; Pred. No. 1,1e-06;
XX	Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 LKFKNKIKKKIKKIGQKI 19
Db	1 LKFKNKIKKKIKKIGQKI 19
XX	
XX	RESULT 2
ID	AAB70670
XX	ID AAB70670 standard; peptide; 21 AA.
XX	
AC	AAB70670;
XX	
DT	15-MAY-2001 (first entry)
XX	
DE	Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:23.
XX	
XX	
KW	Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW	bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW	proliferation inhibitor; gene therapy; <i>Pseudomonas aeruginosa</i> ;
KW	<i>Burkholderia cepacia</i> ; <i>Alcaligenes</i> ; <i>Xanthomonas</i> .
XX	
OS	<i>Oryctolagus cuniculus</i> .
XX	
PN	MO200112668-A1.
XX	
PD	22-FEB-2001.
XX	
PF	18-AUG-2000; 2000MO-US22781.
XX	
PR	18-AUG-1999; 99US-0149886.
XX	
PA	(IOWA) UNIV IOWA RES FOUND.
PA	(RESC) UNIV CALIFORNIA.
XX	
PI	Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX	
DR	WPI: 2001-234911/24.
XX	
PT	New antimicrobial peptides useful as antibiotics for inhibiting growth
PT	and proliferation of microbes, and for treating microbial infections -
XX	
PS	Claim 1; Page 103; 137pp; English.
XX	
CC	AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
CC	AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
CC	SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
CC	derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
CC	antibiotic, antimicrobial and antiviral activities, and can be used as
CC	microbial growth and proliferation inhibitors and in gene therapy. (I)
CC	are useful for inhibiting microbial growth in an environment capable of
CC	sustaining such growth, for inhibiting microbial growth or strain in a
CC	host, and inhibiting the growth of drug-resistant microbial strains such
CC	as <i>Pseudomonas aeruginosa</i> , <i>Burkholderia cepacia</i> , <i>Alcaligenes</i> and
CC	<i>Xanthomonas</i> .
CC	
XX	
SO	Sequence 19 AA;
XX	
XX	
XX	Query Match 100.0%; Score 93; DB 22; Length 19;
XX	Best Local Similarity 100.0%; Pred. No. 1,1e-06;
XX	Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 LKFKNKIKKKIKKIGQKI 19
Db	1 LKFKNKIKKKIKKIGQKI 19
XX	
XX	RESULT 2
ID	AAB70670
XX	ID AAB70670 standard; peptide; 21 AA.
XX	
AC	AAB70670;
XX	
DT	15-MAY-2001 (first entry)
XX	
DE	Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:23.
XX	
XX	
KW	Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW	bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW	proliferation inhibitor; gene therapy; <i>Pseudomonas aeruginosa</i> ;
KW</	

```

CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
CC Xanthomonas.
XX
XX
SQ Sequence 21 AA:

Query Match 100.0%; Score 93; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 1,3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LRRFRNKIKKIKKIGOKI 19
   |||||
Db 1 LKTRNKKIKKIKKIGOKI 19

RESULT 3
AAB70668
ID AAB70668 standard; peptide: 22 AA.
XX
XX AAB70668;
XX AC
XX DT 15-MAY-2001 (first entry)
XX
XX Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:21.
DE
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Oryctolagus cuniculus.
XX OS
XX WO200112668-A1.
XX PN
XX
XX 22-FEB-2001.
XX PD
XX PF 18-AUG-2000; 2000WO-US22781.
XX PR 18-AUG-1999; 99US-0149886.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX PA (RESC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
XX WPI: 2001-234911/24.
XX DR
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
XX PT and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (1), of which
XX CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
XX CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
XX CC derived from the lupine RCAP 18 cathelicidin family peptide. (1) have
XX CC antibiotic, antimicrobial and antiviral activities, and can be used as
XX CC microbial growth and proliferation inhibitors and in gene therapy. (1)
XX CC are useful for inhibiting microbial growth in an environment capable of
XX CC sustaining such growth, for inhibiting microbial growth or strain in a
XX CC host, and inhibiting the growth of drug-resistant microbial strains such
XX CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX CC Xanthomonas.
XX
XX Sequence 22 AA:

Query Match 100.0%; Score 93; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 1,3e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LRRFRNKIKKIKKIGOKI 19
   |||||

```

DB 4 lrkfrnkikexlkkigqki 22

RESULT 4
AAB28486
ID AAB28486 standard; peptide: 32 AA.
XX
AC AAB28486;
XX
DT 02-FEB-2001 (first entry)
XX
DE Rabbit lipopolysaccharide binding protein #1.
XX
KW Rabbit; lipopolysaccharide binding protein; lipopolysaccharide analyte;
KW LPSA; immunoassay; gram-negative bacterium detection; Escherichia coli;
KW Salmonella; Chlamydia; infection.
XX
OS Oryctolagus cuniculus.
XX
PN WO200060354-A1.
XX
PD 12-OCT-2000.
XX
PE 03-APR-2000; 2000MO-EP02869.
XX
PR 07-APR-1999; 99EP-0302711.
XX
PA (UNIL) UNILEVER PLC.
PA (UNIL) UNILEVER NV.
PA (HIND-) HINDUSTAN LEVER LTD.
XX
PI Badley RA, Hughes G, Zak KW;
XX
DR WPI; 2000-679390/66.
XX
PT Immunoassay for detecting lipopolysaccharides (LPS) from gram negative
PT bacteria in a sample, comprises use of LPS-binding protein and an
PT antibody to LPS as first or second binding reagents -
XX
PS Claim 12; Page 13; 40pp; English.
XX
SQ

The present sequence is lipopolysaccharide binding protein which may be used for detecting the presence of a lipopolysaccharide analyte (LPSA) in a sample. Anti-LPSA antibody is also used to detect the LPSA. The immunoassay is useful for detecting the presence of LPSA derived from the cell membrane of a gram-negative bacterium such as Escherichia coli, CC Salmonella and Chlamydia. It provides for assay devices to be constructed which are capable of accurate and reproducible test results, ideally suited for commercial markets such as the clinical or home-testing markets. Furthermore, such immunoassays can be performed quickly and simply, without using complex extracts or multiple assay steps.

Query Match 100.0%; Score 93; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LRRFRNKIKEXLKKIGOKI 19
XXXXXXXXXXXXXXXXXXXX
DB 6 lrkfrnkikexlkkigqki 24

RESULT 5
AAB70665
ID AAB70665 standard; peptide: 37 AA.
XX
AC AAB70665;
XX
DT 15-MAY-2001 (first entry)
XX

DE Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:18.
XX
KW Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
KW bactericidal; antibiotic; antiviral; microbial growth inhibitor;
KW proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
KW Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Oryctolagus cuniculus.
XX
PN WO200112668-A1;
XX
PD 22-FEB-2001;
XX
PE 18-AUG-2000; 2000MO-US22781.
XX
PR 18-AUG-1999; 99US-0149886.
XX
PA (IOWA) UNIV IOWA RES FOUND.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX
DR WPI; 2001-234911/24.
XX
PT New antimicrobial peptides useful as antibiotics for inhibiting growth
PT and proliferation of microbes, and for treating microbial infections -
XX
PS Claim 10; Page 108; 137pp; English.
XX
SQ

AAB70648 to AAB70675 represent antimicrobial peptides (I), of which CC AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine CC SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are CC derived from the lupine RCAP 18 cathelicidin family peptide. (I) have CC antibiotic, antimicrobial and antiviral activities, and can be used as CC microbial growth and proliferation inhibitors and in gene therapy. (I) CC are useful for inhibiting microbial growth in an environment capable of CC sustaining such growth, for inhibiting microbial growth or strain in a CC host, and inhibiting the growth of drug-resistant microbial strains such CC as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and CC Xanthomonas.

Query Match 100.0%; Score 93; DB 22; Length 37;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LRRFRNKIKEXLKKIGOKI 19
XXXXXXXXXXXXXXXXXXXX
DB 6 lrkfrnkikexlkkigqki 24

RESULT 6
AAB51193
ID AAB51193 standard; Peptide: 38 AA.
XX
AC AAB51193;
XX
DT 22-MAR-2001 (first entry)
XX
DE E. coli AMP gene CAP18 amino acid sequence.
XX
KW Escherichia coli; E. coli; AMP gene; anti-microbial peptide;
KW screening; preservation; food; feed; paint formulation; detergent;
KW cosmetic; medical device; prosthetic implant; disinfectant;
KW microbial infection; tumour.
XX
OS Escherichia coli.
XX
PN WO200073433-A1.
XX
PD 07-DEC-2000.

XX 29-MAY-2000; 2000WO-DK00287.
 PF of nucleotide sequences (NT) linked to an inducible promoter, to express
 XX a peptide (P) (an enzyme or mature (P) of less than 100 amino acids
 PR 31-MAY-1999; 99DK-0000766.
 XX optionally linked to a signal (P)), transformed into host cells and
 PA (NOVO) NOVO NORDISK AS.
 XX
 PI Kristensen H;
 XX
 DR WPI: 2001-070965/08.
 XX
 PT Screening for nucleotide sequences encoding antimicrobial peptides by
 PT generating peptide libraries in microorganisms, inducing expression of
 PT peptides, selecting cells and recovering nucleotide sequences from
 PT cells
 XX
 PS Example 1: Page 34; 59pp; English.
 XX
 CC The present invention describes a plasmid which is ligated with a pool
 CC of nucleotide sequences (NT) linked to an inducible promoter, to express
 CC a peptide (P) (an enzyme or mature (P) of less than 100 amino acids
 CC optionally linked to a signal (P)), transformed into host cells and
 CC cultured in presence of an inducer to induce expression of the NT. A
 CC method of screening (I) a pool of nucleotide sequences to select a
 CC nucleotide sequence encoding a peptide, comprises: (a) ligating a plasmid
 CC with the pool of NT; (b) transforming host cells which are sensitive to
 CC the peptide with the ligated plasmids; (c) screening the transformed
 CC cells to select viable cells; (d) cultivating the viable cells in the
 CC presence of an inducer to induce expression of NT; (e) selecting cells
 CC according to the effect of the inducer on cell growth; and (f) recovering
 CC NT encoding the peptide from the selected cells. (f) is useful for
 CC screening a pool of nucleotide sequences to select a nucleotide sequence
 CC encoding a peptide which is an antimicrobial peptide or an antimicrobial
 CC enzyme active on bacteria and for finding and preparing a composition for
 CC treatment of human or animal. The antimicrobial peptide obtained using
 CC (I) may be employed in preservation of e.g. food/feed, paint
 CC formulations, detergents, cosmetics, medical devices such as prosthetic
 CC implants and also to disinfect and/or kill microbial cells on an object
 CC e.g. as an disinfectant for the treatment of biofilm. The peptides are
 CC useful for treating microbial infections and/or tumours. Peptides with
 CC improved bio-activity can be developed using (I). The peptides have no
 CC negative effect on normal mammalian and/or eukaryotic cells. The present
 CC sequence represents an AMP (antimicrobial peptide) gene amino acid
 CC sequence, which is used in an example from the present invention.
 CC
 XX
 SQ Sequence 38 AA:
 QY 1 LRRFRNKKIKKKIKGOKI 19
 Db 7 lrkfrnkikkkikkgqki 25
 Query Match 100.0%; Score 93; DB 22; Length 38;
 Best Local Similarity 100.0%; Pred. No. 2.2e-06;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FH Key Location/Qualifiers
 FT Misc-difference 23 /label= Asp, Lys
 FT Misc-difference 26 /label= Gln, Ile
 FT Misc-difference 27 /label= Gly, Gln
 FT
 PN W09402589-A.
 XX
 PD 03-FEB-1994.
 XX
 PF 15-JUL-1993; 93WO-US06731.
 XX
 PR 17-JUL-1992; 92US-0916761.
 PR 17-JUL-1992; 92US-0916765.
 PA (PANO-) PANORAMA RES INC.
 XX
 PI Hirata M, Larrick JW, Wright SC;
 XX
 DR WPI: 1994-048847/06.
 XX
 PT Sequences encoding mammalian cationic antibacterial proteins
 PT are homologous to human and rabbit CAP18 sequences and have
 PT lipo:polysaccharide binding and anti-coagulation activity
 XX
 PS Disclosure: Page 59; 112pp; English.
 XX
 CC The sequence is that of a human cationic antibacterial protein CAP18
 CC N-terminal fragment obd. from granulocytes as sequenced. The
 CC sequence corresponds to the N-terminus of the C-terminal RNIP fragment
 CC of rabbit CAP18. The fragment is capable of binding to lipopolysaccharide
 CC and inhibiting LPS-mediated activation of macrophage, as well as
 CC interfering with the clotting cascade to inhibit coagulation in
 CC conditions of disseminated intravascular coagulation. The polypeptides
 CC can also be used to attenuate, inhibit or prevent LPS-associated
 CC conditions, e.g. sepsis, autoimmune disorders, inflammation, etc.
 CC See also AAR4567-81.
 CC
 XX
 SQ Sequence 29 AA:
 QY 1 LRRFRNKKIKKKIKGOKI 19
 Db 6 lrkfrnkikkkikkgqxi 24
 Query Match 93.5%; Score 87; DB 15; Length 29;
 Best Local Similarity 94.7%; Pred. No. 1.2e-05;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 8
 AAB07903
 ID AAB07903 standard; Protein: 29 AA.
 XX
 AC AAB07903;
 XX
 DT 14-NOV-2000 (first entry)
 DE Putative N-terminal of a human cationic protein designated CAP18.
 XX
 KW Human; cationic protein; lipopolysaccharide binding; anticoagulant;
 KW CAP18; reactive nitrogen inhibitory protein; RNIP; gram negative sepsis;
 KW coagulation-related disorder; disseminated intravascular coagulation.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 23 /label= Asp, Lys
 FT Misc-difference 26 /label= Gln, Ile
 FT Misc-difference 26 /label= Gln, Ile


```

XX 15-MAY-2001 (first entry)
DT Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:25.
XX
XX
DE Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
OS Oryctolagus cuniculus.
XX WO200112668-A1.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000WO-US22781.
XX
XX 18-AUG-1999; 99US-0149886.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
XX AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
XX SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
XX derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
XX antibiotic, antimicrobial and antiviral activities, and can be used as
XX microbial growth and proliferation inhibitors and in gene therapy. (I)
XX are useful for inhibiting microbial growth in an environment capable of
XX sustaining such growth, for inhibiting microbial growth or strain in a
XX host, and inhibiting the growth of drug-resistant microbial strains such
XX as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX Xanthomonas.
XX
SQ Sequence 15 AA:

Query Match 66.7%; Score 62; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 KIKKKIKKIGQKI 19
   |||||||
Db 1 KIKKKIKKIGQKI 13

RESULT 14
AAB70673
ID AAB70673 standard; peptide; 17 AA.
XX
XX AAB70673;
AC
XX
XX 15-MAY-2001 (first entry)
DT
XX
XX Lupine RCAP 18 cathelicidin derived antimicrobial peptide SEQ ID NO:26.
DE
XX
XX Ovine; SMAP29; lupine; RCAP 18; cathelicidin; antimicrobial;
XX bactericidal; antibiotic; antiviral; microbial growth inhibitor;
XX proliferation inhibitor; gene therapy; Pseudomonas aeruginosa;
XX Burkholderia cepacia; Alcaligenes; Xanthomonas.
XX
XX Oryctolagus cuniculus.
XX

```

```

PN WO200112668-A1.
XX
XX 22-FEB-2001.
PD
XX
XX 18-AUG-2000; 2000WO-US22781.
PF
XX
XX 18-AUG-1999; 99US-0149886.
PR
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Tack BE, McCray P, Welsh M, Travis SM, Lehrer R;
XX WPI; 2001-234911/24.
XX
XX New antimicrobial peptides useful as antibiotics for inhibiting growth
XX and proliferation of microbes, and for treating microbial infections -
XX
XX Claim 1; Page 103; 137pp; English.
XX
XX AAB70648 to AAB70675 represent antimicrobial peptides (I), of which
XX AAB70648 to AAB70664, AAB70674 and AAB70675 are derived from the ovine
XX SMAP 29 cathelicidin family peptide, and AAB70665 to AAB70673 are
XX derived from the lupine RCAP 18 cathelicidin family peptide. (I) have
XX antibiotic, antimicrobial and antiviral activities, and can be used as
XX microbial growth and proliferation inhibitors and in gene therapy. (I)
XX are useful for inhibiting microbial growth in an environment capable of
XX sustaining such growth, for inhibiting microbial growth or strain in a
XX host, and inhibiting the growth of drug-resistant microbial strains such
XX as Pseudomonas aeruginosa, Burkholderia cepacia, Alcaligenes and
XX Xanthomonas.
XX
SQ Sequence 17 AA:

Query Match 66.7%; Score 62; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 KIKKKIKKIGQKI 19
   |||||||
Db 1 KIKKKIKKIGQKI 13

RESULT 15
AAG66421
ID AAG66421 standard; peptide; 36 AA.
XX
XX AAG66421;
AC
XX
XX 23-OCT-2001 (first entry)
DT
XX
XX Mouse cathelin peptide.
DE
XX
XX Mouse; cathelin; immunomodulator; polycationic polymer; vaccine;
XX immune response.
XX
XX Mus sp.
OS
XX
XX Key Location/Qualifiers
FT Misc-difference 22 /note= "Residue given as "O" in specification"
FT
XX
XX WO200154720-A1.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 05-JAN-2001; 2001WO-EP00087.
PF
XX
XX 28-JAN-2000; 2000AT-0000129.
PR
XX
XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX

```

PI Lingnau K, Mattner F, Schmidt W, Birnstiel M, Buschle M;
 XX
 DR WPI; 2001-536419/59.
 XX
 PT Pharmaceutical composition useful for inducing immune response
 PT comprises antigen, immunogenic oligodeoxynucleotide containing
 PT cytosine-guanine dinucleotide motifs and polycationic polymer -
 XX
 PS Disclosure; Page 12; 39pp; English.
 XX
 CC The present invention relates to a pharmaceutical composition which
 CC comprises an antigen, an immunogenic oligodeoxynucleotide containing
 CC cytosine-guanine dinucleotide (CpG) motifs (CpG-ODN) and a polycationic
 CC polymer. The composition is useful for making a vaccine to induce potent
 CC immune responses, or to decrease or ablate undesired immune responses.
 CC The present sequence is mouse cathelin peptide, which was used as a
 CC polycationic compound in the present invention.
 XX
 SQ Sequence 36 AA;

Query Match 62.4%; Score 58; DB 22; Length 36;
 Best Local Similarity 73.7%; Pred. No. 0.18;
 Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Oy 1 LKKFRKKIKKKIKKKI 19
 ||| | | | | | | | |
 Db 6 lrkgeklgeklklyxki 24

Search completed: July 12, 2002, 08:00:44
 Job time: 451 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:55:09 ; Search time 75.52 Seconds
(without alignments)
6.145 Million cell updates/sec

Title: US-09-642-744B-21

Perfect score: 93

Sequence: 1 LKFRNKIKKIKGOKI 19

Scoring table: BLOSUMP62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	171	1	US-08-313-681A-4
2	93	100.0	171	3	US-09-322-911-4
3	87	93.5	29	1	US-08-313-681A-7
4	87	93.5	29	3	US-09-322-911-7
5	64	68.8	33	5	PCT-US95-12080-4
6	55	59.1	16	1	US-08-313-681A-11
7	55	59.1	16	3	US-09-322-911-11
8	47	50.5	190	1	US-08-393-985-25
9	47	50.5	263	1	US-08-393-985-23
10	47	50.5	285	1	US-08-393-985-2
11	47	50.5	288	1	US-08-337-602-4
12	47	50.5	288	3	US-08-556-135-4
13	47	50.5	288	3	US-08-819-286-3
14	46	49.5	27	1	US-08-231-730A-15
15	46	49.5	27	1	US-08-427-001C-15
16	46	49.5	27	1	US-08-457-798-15
17	46	49.5	27	1	US-08-457-798-15
18	46	49.5	27	2	US-08-505-486-15
19	46	49.5	27	2	US-08-505-486-62
20	46	49.5	27	3	US-08-689-489C-15
21	46	49.5	27	3	US-08-801-028-15
22	46	49.5	27	3	US-08-801-028-62
23	46	49.5	27	3	US-09-340-154-15
24	46	49.5	27	3	US-09-340-154-62
25	46	49.5	27	4	US-09-232-802A-15
26	46	49.5	27	5	PCT-US94-06176-15
27	46	49.5	27	5	PCT-US94-12550-15

28	46	49.5	27	5	PCT-US95-04335-15	Sequence 15, Appl
29	46	49.5	27	5	PCT-US95-04718-15	Sequence 15, Appl
30	46	49.5	27	5	PCT-US95-09338-15	Sequence 15, Appl
31	46	49.5	27	5	PCT-US95-09338-62	Sequence 62, Appl
32	46	49.5	27	5	PCT-US95-09339-15	Sequence 15, Appl
33	46	49.5	27	5	PCT-US95-09339-62	Sequence 62, Appl
34	46	49.5	32	1	US-08-231-730A-16	Sequence 16, Appl
35	46	49.5	32	1	US-08-427-001C-16	Sequence 16, Appl
36	46	49.5	32	1	US-08-457-798-16	Sequence 16, Appl
37	46	49.5	32	1	US-08-457-171-16	Sequence 16, Appl
38	46	49.5	32	2	US-08-505-486-16	Sequence 16, Appl
39	46	49.5	32	3	US-08-689-489C-16	Sequence 16, Appl
40	46	49.5	32	3	US-08-801-028-16	Sequence 16, Appl
41	46	49.5	32	3	US-09-340-154-16	Sequence 16, Appl
42	46	49.5	32	4	US-09-232-802A-16	Sequence 16, Appl
43	46	49.5	32	5	PCT-US94-06176-16	Sequence 16, Appl
44	46	49.5	32	5	PCT-US94-12550-16	Sequence 16, Appl
45	46	49.5	32	5	PCT-US95-04335-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-313-681A-4
; Sequence 4, Application US/08313681A
; Patent No. 5618675
; GENERAL INFORMATION:
; APPLICANT: Larrick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Mishimasa
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,681A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 15325-9-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-313-681A-4
Query Match 100.0%, Score 93; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 9.9e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LKFRNKIKKIKGOKI 19
|||||

TITLE OF INVENTION: Human Cationic Proteins Having
TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/322,911
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/691,280
FILING DATE: August 1, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,761
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/916,765
FILING DATE: July 17, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06731
FILING DATE: July 15, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/313,681
FILING DATE: September 27, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 15325-000920
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 23
OTHER INFORMATION: /note= "Xaa is Asp or Lys"
FEATURE:
NAME/KEY: Region
LOCATION: 26
OTHER INFORMATION: /note= "Xaa is a Gln or Ile"
FEATURE:
NAME/KEY: Region
LOCATION: 27
OTHER INFORMATION: /note= "Xaa is a Gly or Gln"
US-09-322-911-7

Query Match 93.5%; Score 87; DB 3; Length 29;
Best Local Similarity 94.7%; Pred. No. 1.2e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LRFKRNKIKKIKKIGOKI 19
|||||
Db 6 LRFKRNKIKKIKKIGOKI 24
|||||
RESULT 5

PCT-US95-12080-4
Sequence 4, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporaton
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-12080-4

Query Match 68.8%; Score 64; DB 5; Length 33;
Best Local Similarity 78.9%; Pred. No. 0.016;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LRFKRNKIKKIKKIGOKI 19
|||||
Db 3 LRFKRNKIKKIKKIGOKI 21

RESULT 6
US-08-313-681A-11
Sequence 11, Application US/08313681A
Patent No. 5618675
GENERAL INFORMATION:
APPLICANT: Larrick, James W.
APPLICANT: Wright, Susan C.
APPLICANT: Hirata, Mshimasa
TITLE OF INVENTION: Human Cationic Proteins Having
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,681A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Heslin, James M.
REGISTRATION NUMBER: 29,541


```

; REFERENCE/DOCKET NUMBER: 15325-9-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-681A-11

Query Match          59.1%; Score 55; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKFRNKIKRK 11
DB 6 LKFRNKIKRK 16

RESULT 7
US-09-322-911-11
; Sequence 11, Application US/09322911
; Patent No. 6103888
; GENERAL INFORMATION:
; APPLICANT: Larick, James W.
; APPLICANT: Wright, Susan C.
; APPLICANT: Hirata, Nobumasa
; APPLICANT: Balint, Robert F.
; TITLE OF INVENTION: Human Cationic Proteins Having
; TITLE OF INVENTION: Lipopolysaccharide Binding and Anti-Coagulant Activity
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/322,911
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/691,280
; FILING DATE: August 1, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,761
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/916,765
; FILING DATE: July 17, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06731
; FILING DATE: July 15, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/313,681
; FILING DATE: September 27, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 15325-000920
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
```

```

; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-322-911-11

Query Match          59.1%; Score 55; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKFRNKIKRK 11
DB 6 LKFRNKIKRK 16

RESULT 8
US-08-393-985-25
; Sequence 25, Application US/08393985
; Patent No. 5693476
; GENERAL INFORMATION:
; APPLICANT: Scheller, Richard H.
; TITLE OF INVENTION: Methods and Compositions for Modulation
; TITLE OF INVENTION: of Vesicular Release
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,985
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8600-0152
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 190 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: amino acids 4-193 of syn1A
US-08-393-985-25

Query Match          50.5%; Score 47; DB 1; Length 190;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKFRNKIKRKIKGKI 19
DB 79 IKTKANKVSKLSIESOI 97

RESULT 9
US-08-393-985-23
```

Sequence 23, Application US/08393985
Patent No. 5693476
GENERAL INFORMATION:
APPLICANT: Scheller, Richard H.
TITLE OF INVENTION: Methods and Compositions for Modulation
TITLE OF INVENTION: of Vesicular Release
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,985
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0152
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-393-985-23

Query Match 50.5%; Score 47; DB 1; Length 263;
Best Local Similarity 47.4%; Pred. No. 21;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKFRNKIKKIKKIGOKI 19
Db 79 IKKTANKVRSKLSIESQSI 97

RESULT 10
US-08-393-985-2
Sequence 2, Application US/08393985
Patent No. 5693476
GENERAL INFORMATION:
APPLICANT: Scheller, Richard H.
TITLE OF INVENTION: Methods and Compositions for Modulation
TITLE OF INVENTION: of Vesicular Release
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,985
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0152
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-393-985-2

Query Match 50.5%; Score 47; DB 1; Length 285;
Best Local Similarity 47.4%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKFRNKIKKIKKIGOKI 19
Db 79 IKKTANKVRSKLSIESQSI 97

RESULT 11
US-08-337-602-4
Sequence 4, Application US/08337602
Patent No. 5623051
GENERAL INFORMATION:
APPLICANT: Catterall, William A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
TITLE OF INVENTION: FOR PRESYNAPTIC CALCIUM CHANNEL BLOCKERS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98042-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/337,602
FILING DATE: 10-NOV-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.602
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANDBERRY
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-337-602-4

Query Match 50.5%; Score 47; DB 1; Length 288;
Best Local Similarity 47.4%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 LKFRNKIKKIKKIGOKI 19

Db 82 IKTKANKVRSKLKSIQSI 100

RESULT 12

US-08-558-135-4
Sequence 4, Application US/08558135
Patent No. 6090631
GENERAL INFORMATION:
APPLICANT: Catterall, William A.
APPLICANT: Sheng, Zu-Hang
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/558,135
FILING DATE: 13-NOV-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.602C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-558-135-4

Query Match 50.5%; Score 47; DB 3; Length 288;
Best Local Similarity 47.4%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKKFRNKIKKIKIGQKI 19
Db 82 IKTKANKVRSKLKSIQSI 100

RESULT 13
US-08-819-286-3
Sequence 3, Application US/08819286
Patent No. 6169074
GENERAL INFORMATION:
APPLICANT: Montal, Mauricio
TITLE OF INVENTION: PEPTIDE INHIBITORS OF
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,286
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,599
FILING DATE: 18-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Taylor, Stacy L.
REGISTRATION NUMBER: 34,842
REFERENCE/DOCKET NUMBER: 07349/005001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 288 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-819-286-3

Query Match 50.5%; Score 47; DB 4; Length 288;
Best Local Similarity 47.4%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LKKFRNKIKKIKIGQKI 19
Db 82 IKTKANKVRSKLKSIQSI 100

RESULT 14
US-08-231-730A-15
Sequence 15, Application US/08231730A
Patent No. 5561107
GENERAL INFORMATION:
APPLICANT: JAYNES, JESSE M.
APPLICANT: JULIAN, GORDON R.
TITLE OF INVENTION: METHOD OF ENHANCING WOUND HEALING BY STIMULATING FIBROBLAST
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: STEVEN J. HULTQUIST
ADDRESS: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
STREET: 200 PARK DRIVE, SUITE 210
STREET: P.O. BOX 14329
CITY: RESEARCH TRIANGLE PARK
STATE: NORTH CAROLINA
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: APPLE MACINTOSH
OPERATING SYSTEM: MACINTOSH
SOFTWARE: M.S. WORD 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231,730A
FILING DATE: 04-20-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-08-94
APPLICATION NUMBER: 08/039,620
FILING DATE: 06-04-93
APPLICATION NUMBER: 08/148,491
FILING DATE: 11-08-93
APPLICATION NUMBER: 08/148,889
FILING DATE: 11-08-93
ATTORNEY/AGENT INFORMATION:
NAME: HULTQUIST, STEVEN J.

REGISTRATION NUMBER: 28021
REFERENCE/DOCKET NUMBER: 4013-106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)990-9531
TELEFAX: (919)990-9532
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-231-730A-15

Query Match 49.5%; Score 46; DB 1; Length 27;
Best Local Similarity 38.9%; Pred. No. 3.3;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 KRFNKKIKKKIKGKI 19
:| | :| | :| |
Db 3 KKFVKKVAKKKVAKKV 20

RESULT 15
US-08-427-001C-15
Sequence 15, Application US/08427001C
Patent No. 5717064
GENERAL INFORMATION:
APPLICANT: JULIAN, GORDON R.
TITLE OF INVENTION: METHYLATED LYSINE-RICH LYTIC PEPTIDES,
TITLE OF INVENTION: AND METHOD OF MAKING THE SAME BY REDUCTIVE ALKYLATION
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/427,001C
FILING DATE: 24-APR-95
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/148,889
FILING DATE: 08-NOV-93
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: WALKER, BARBARA W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 2093-105A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
US-08-427-001C-15

Query Match 49.5%; Score 46; DB 1; Length 27;
Best Local Similarity 38.9%; Pred. No. 3.3;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
QY 2 KRFNKKIKKKIKGKI 19
:| | :| | :| |
Db 3 KKFVKKVAKKKVAKKV 20

Search completed: July 12, 2002, 07:55:09
Job time: 116 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 07:56:57 ; Search time 95.47 Seconds

(without alignments)
19.123 Million cell updates/sec

Title: US-09-642-744B-21

Perfect score: 93

Sequence: 1 LRFRNKIKKKIKGKI 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR_71:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	100.0	171	2 J01171	18K lipopolysaccha
2	64	68.8	173	2 S70521	cathelin-related p
3	55	58.1	324	2 T07982	probable choline-p
4	55	59.1	326	2 T07983	choline-phosphate
5	52	55.9	329	2 T07981	probable choline-p
6	52	55.9	331	2 T07980	probable choline-p
7	49	52.7	142	2 E70356	hypothetical prote
8	49	52.7	332	2 H84730	probable phosphol
9	49	52.7	1163	1 GNMVY8	genome polypeptid
10	49	52.7	3411	1 GNMVY8	genome polypeptid
11	48	51.6	152	2 C72227	conserved hypotet
12	48	51.6	293	2 T04951	hypothetical prote
13	48	51.6	497	2 E90061	hypothetical prote
14	47.5	51.1	274	2 F84426	conserved hypotet
15	47	50.5	174	2 F84426	hypothetical prote
16	47	50.5	259	2 G01485	synthaxin - human
17	47	50.5	288	2 J01615	neuron-specific an
18	47	50.5	298	2 A38141	neuronal cell memb
19	47	50.5	507	2 C82901	conserved hypotet
20	47	50.5	507	2 C82901	antibacterial prot
21	46	49.5	166	2 A41731	polypeptide deform
22	46	49.5	169	2 A82373	synthaxin 1B - rat
23	46	49.5	288	2 J00136	synthaxin 1B - rat
24	46	49.5	288	2 B48213	hypothetical prote
25	46	49.5	362	2 T22204	hypothetical prote
26	46	49.5	519	2 T51496	hypothetical prote
27	46	49.5	1236	2 T18459	hypothetical prote
28	45	48.4	189	2 C64469	hypothetical prote
29	45	48.4	443	2 A82872	ATP-dependent RNA

30	45	48.4	486	2 S74319	RTG3 protein - yea
31	45	48.4	522	2 A95059	hypothetical prote
32	45	48.4	522	2 H97927	type 1 site-specif
33	45	48.4	573	2 D90202	methionyl-tRNA syn
34	45	48.4	663	2 D97047	DNA ligase (NAD de
35	45	48.4	1631	1 SA20K1	major mezozoite su
36	45	48.4	1639	2 S05603	major mezozoite su
37	45	48.4	1640	2 A24594	probable major sur
38	44.5	47.8	121	2 H70471	conserved hypotet
39	44.5	47.8	399	2 D86322	hypothetical prote
40	44	47.3	109	2 G64379	hypothetical prote
41	44	47.3	119	2 C90351	hypothetical prote
42	44	47.3	175	2 A64330	hypothetical prote
43	44	47.3	188	2 H97293	hypoxanthine-guan
44	44	47.3	182	2 E71157	probable CDP-alcoh
45	44	47.3	241	2 G86355	hypothetical prote

ALIGNMENTS

```

RESULT 1
J01171
18K lipopolysaccharide-binding protein precursor - rabbit
N:Alternate names: 18K cationic protein
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 06-Dec-1996
C:Accession: J01171; PS0226
R:Larriek, J.W.; Morgan, J.G.; Palings, I.; Hirata, M.; Yen, M.H.
Biochem. Biophys. Res. Commun. 179, 170-175, 1991
A:Title: Complementary DNA sequence of rabbit CAP18-a unique lipopolysaccharide bindi
A:Reference number: J01171; MUID:91354246
A:Accession: J01171
A:Molecule type: mRNA
A:Residues: 1-171 <LAR>
A:Experimental source: bone marrow
A:Accession: PS0226
A:Molecule type: protein
A:Residues: 135-159, 'GTGQL' <LAR>
A:Note: 157-Asp was also found
C:Superfamily: cathelin; cystatin homology
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-171/Product: 18K lipopolysaccharide-binding protein #status predicted <MAT>

Query Match 100.0%; Score 93; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LRFRNKIKKKIKGKI 19
Db 140 LRFRNKIKKKIKGKI 158

RESULT 2
S70521
cathelin-related protein precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S70521
R:Popstev, A.E.; Zinovjeva, M.V.; Visser, J.W.M.; Zijlman, J.M.J.M.; Fibbe, W.E.; B
FEBS Lett. 391, 5-8, 1996
A:Title: A novel murine cathelin-like protein expressed in bone marrow.
A:Reference number: S70521; MUID:96326596
A:Accession: S70521
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-173 <POP>
A:Cross-references: EMBL:X94333; NID:g1177533; PIDN:CAA64078.1; PID:g1177534
C:Superfamily: cathelin; cystatin homology
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-144/Domain: propeptide #status predicted <PRO>
F:145-173/Product: cathelin-related protein #status predicted <MAT>

```

Query Match 68.8%; Score 64; DB 2; Length 173;
Best Local Similarity 78.9%; Pred. No. 0.18;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LKFRNKIKKELKKIGKI 19
DB 142 LKKGEXKIGKELKKIGKI 160

RESULT 3
T07982
probable choline-phosphate cytidylyltransferase (EC 2.7.7.15) (clone CCT3) - rape
N:Alternate names: CTP:phosphocholine cytidylyltransferase
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T07982
R:Nishida, I.; Swinhoe, R.; Slabas, A.R.; Murata, N.
Plant Mol. Biol. 31, 205-211, 1996
A:Title: Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase cDNAs by comp
A:Reference number: 216266; MUID:96343925
A:Accession: T07982
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-324 <NIS>
A:Cross-references: EMBL:D63167; NID:d1049575; PIDN:BAA09643.1; PID:d1010288
A:Experimental source: cv. Jet Nurfi; roots
A:Gene: CCT
C:Function:
A:Description: catalyzes the synthesis of CDP-choline and PPI from CTP and phosphocholin
C:Keywords: nucleotidyltransferase

Query Match 59.1%; Score 55; DB 2; Length 324;
Best Local Similarity 47.4%; Pred. No. 5.1;
Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKFRNKIKKELKKIGKI 19
DB 197 LKKLOEKVKKEQEKVGKGI 215

RESULT 4
T07983
choline-phosphate cytidylyltransferase (EC 2.7.7.15) (clone CCT4) - rape
N:Alternate names: CTP:phosphocholine cytidylyltransferase
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T07983
R:Nishida, I.; Swinhoe, R.; Slabas, A.R.; Murata, N.
Plant Mol. Biol. 31, 205-211, 1996
A:Title: Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase cDNAs by comp
A:Reference number: 216266; MUID:96343925
A:Accession: T07983
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-326 <NIS>
A:Cross-references: EMBL:D63168; NID:d1049576; PIDN:BAA09644.1; PID:d1010289
A:Experimental source: cv. Jet Nurfi; roots
A:Gene: CCT
C:Function:
A:Description: catalyzes the synthesis of CDP-choline and PPI from CTP and phosphocholin
C:Keywords: nucleotidyltransferase

Query Match 59.1%; Score 55; DB 2; Length 326;
Best Local Similarity 47.4%; Pred. No. 5.1;
Matches 9; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKFRNKIKKELKKIGKI 19

DB 199 LKKLOEKVKKEQEKVGKGI 217

RESULT 5
T07981
probable choline-phosphate cytidylyltransferase (EC 2.7.7.15) (clone CCT1) - rape
N:Alternate names: CTP:phosphocholine cytidylyltransferase
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T07981
R:Nishida, I.; Swinhoe, R.; Slabas, A.R.; Murata, N.
Plant Mol. Biol. 31, 205-211, 1996
A:Title: Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase cDNAs by c
A:Reference number: 216266; MUID:96343925
A:Accession: T07981
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-329 <NIS>
A:Cross-references: EMBL:D58404; NID:d1044869; PIDN:BAA09571.1; PID:d1010214
A:Experimental source: cv. Jet Nurfi; roots
A:Gene: CCT
C:Function:
A:Description: catalyzes the synthesis of CDP-choline and PPI from CTP and phosphochoc
C:Keywords: nucleotidyltransferase

Query Match 55.9%; Score 52; DB 2; Length 329;
Best Local Similarity 42.1%; Pred. No. 13;
Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKFRNKIKKELKKIGKI 19
DB 206 LKKLOEKVKKEQEKVGKGI 224

RESULT 6
T07980
probable choline-phosphate cytidylyltransferase (EC 2.7.7.15) (clone CCT2) - rape
N:Alternate names: CTP:phosphocholine cytidylyltransferase
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T07980
R:Nishida, I.; Swinhoe, R.; Slabas, A.R.; Murata, N.
Plant Mol. Biol. 31, 205-211, 1996
A:Title: Cloning of Brassica napus CTP:phosphocholine cytidylyltransferase cDNAs by c
A:Reference number: 216266; MUID:96343925
A:Accession: T07980
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-331 <NIS>
A:Cross-references: EMBL:D63166; NID:d1049574; PIDN:BAA09642.1; PID:d1010287
A:Experimental source: cv. Jet Nurfi; roots
A:Gene: CCT
C:Function:
A:Description: catalyzes the synthesis of CDP-choline and PPI from CTP and phosphochoc
C:Keywords: nucleotidyltransferase

Query Match 55.9%; Score 52; DB 2; Length 331;
Best Local Similarity 42.1%; Pred. No. 13;
Matches 8; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 LKFRNKIKKELKKIGKI 19
DB 206 LKKLOEKVKKEQEKVGKGI 224

RESULT 7
E70356
hypothetical protein aq_635 - Aquifex aeolicus

C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 22-Jun-1999
C:Accession: E70356
R:Deckerl, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666
A:Accession: E70356
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-142 <AQF>
A:Cross-References: GB:AE000699; NID:g2983238; PIDN:AA06842.1; PID:g2983246; GB:AE00065
A:Experimental source: strain VFS
C:Genetics:
A:Gene: ag 635
C:Superfamily: protein-tyrosine-phosphatase, low molecular weight

Query Match 52.7%; Score 49; DB 2; Length 142;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 LRRFRNKIKKIKKI 15
DB 115 LRRIRDKIEELKKL 129

RESULT 8
H84730
Probable phospholipid cytidyltransferase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84730
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Motilal, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487
A:Accession: H84730
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-332 <STO>
A:Cross-References: GB:AE002093; NID:g3831468; PIDN:AA06950.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g32260
A:Map position: 2

Query Match 52.7%; Score 49; DB 2; Length 332;
Best Local Similarity 36.8%; Pred. No. 32;
Matches 7; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

OY 1 LRRFRNKIKKIKKI 19
DB 207 LKKLGRVKEQOERVEKI 225

RESULT 9
GMVY8
genome polypotein - yellow fever virus (strain 1899/81) (fragment)
N:Contains: amino end of nonstructural protein NS1; capsid protein C; envelope protein E
C:Species: yellow fever virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001
C:Accession: J00374
R:Ballinger-Crabtree, M.E.; Miller, B.R.
J. Gen. Virol. 71, 2115-2121, 1990
A:title: Partial nucleotide sequence of South American yellow fever virus strain 1899/81
A:Reference number: J00374; MUID:91011358
A:Accession: J00374
A:Molecule type: genomic RNA
A:Residues: 1-1163 <BAL>

A:Cross-References: GB:D14458; GB:D00739; NID:g2222777; PIDN:BA03355.1; PID:g2222778
C:Superfamily: yellow fever virus genome polypotein
C:Keywords: capsid protein; envelope protein; glycoprotein; nonstructural protein; nu
F:2-121/Product: capsid protein C #status predicted <CAP>
F:106-122/Domain: transmembrane #status predicted <TM1>
F:122-285/Product: envelope protein M #status predicted <ERM>
F:251-267/Domain: transmembrane #status predicted <TM2>
F:271-287/Domain: transmembrane #status predicted <TM3>
F:286-778/Product: major envelope protein E #status predicted <ENP>
F:733-753/Domain: transmembrane #status predicted <TM4>
F:762-778/Domain: transmembrane #status predicted <TM5>
F:779-1163/Product: nonstructural protein NS1 (fragment) #status predicted <NS1>
F:1068-1075/Region: nucleotide-binding motif A (P-loop)
F:1133-1151/Domain: transmembrane #status predicted <TM6>
F:134,150,172,266,554,755,908,986/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 52.7%; Score 49; DB 1; Length 1163;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 LRRFRNKIKKIKKI 18
DB 20 VRSLNKKIKKIKKI 37

RESULT 10
GMVY
genome polypotein - yellow fever virus (strain 170)
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstruct
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: yellow fever virus
C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 19-Jan-2001
C:Accession: A03914
R:Rice, C.M.; Lenches, E.M.; Eddy, S.R.; Shin, S.J.; Sheets, R.L.; Strauss, J.H.
Science 229, 726-733, 1985
A:title: Nucleotide sequence of yellow fever virus: implications for flavivirus gene
A:Reference number: A03914; MUID:85272570
A:Accession: A03914
A:Molecule type: genomic RNA
A:Residues: 1-3411 <RIC>
A:Cross-References: GB:X03700; GB:K02749; NID:g59338; PIDN:CAA27332.1; PID:g59339
C:Superfamily: yellow fever virus genome polypotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protei
F:2-210/Product: capsid protein C #status predicted <CPC>
F:211-285/Product: envelope protein M #status predicted <EPM>
F:249-269/Domain: transmembrane #status predicted <TM1>
F:271-285/Domain: transmembrane #status predicted <TM2>
F:286-778/Product: major envelope protein E #status predicted <MEE>
F:740-753/Domain: transmembrane #status predicted <TM3>
F:755-778/Domain: transmembrane #status predicted <TM4>
F:779-1187/Product: nonstructural protein NS1 #status predicted <NS1>
F:1159-1180/Domain: transmembrane #status predicted <TM5>
F:1188-1354/Product: nonstructural protein NS2a #status predicted <N2A>
F:1355-1484/Product: nonstructural protein NS2b #status predicted <N2B>
F:1485-2107/Product: nonstructural protein NS3 #status predicted <NS3>
F:1682-1689/Region: nucleotide-binding motif A (P-loop)
F:1769-1774/Region: nucleotide-binding motif B
F:1773-1776/Region: DEAH motif
F:2108-2394/Product: nonstructural protein NS4a #status predicted <N4A>
F:2395-2506/Product: nonstructural protein NS4b #status predicted <N4B>
F:2507-3411/Product: nonstructural protein NS5 #status predicted <NS5>
F:134,150,172,266,554,755,908,986,1796,2062,2320,2346,2408,2467,2720,2734,2740/Bindin

Query Match 52.7%; Score 49; DB 1; Length 3411;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 LRRFRNKIKKIKKI 18
DB 20 VRSLNKKIKKIKKI 37

RESULT 11

genome polyprotein - yellow fever virus (strain Pasteur 17D-204)
N:Contains: capsid protein C; envelope protein M; major envelope protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: yellow fever virus
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 19-Jan-2001
C:Accession: S07757
R:Dupuy, A.; Despres, P.; Cahour, A.; Girard, M.; Bouloy, M.
Nucleic Acids Res. 17, 3989, 1989
A:Title: Nucleotide sequence comparison of the genome of two 17D-204 yellow fever vaccine strains
A:Reference number: S07757; MUID:89282413
A:Accession: S07757
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-3411 <DUP>
A:Cross-references: EMBL:X15062; NID:962289; PIDN:CAB37419.1; PID:94456986
A>Note: The nucleotide sequence was submitted to the EMBL Data Library, April 1989, in C
C:Superfamily: yellow fever virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; nonstructural protein;
F:2-310/Product: capsid protein C #status predicted <CPC>
F:105-125/Product: envelope protein M #status predicted <EPM>
F:211-285/Product: envelope protein M #status predicted <EPM>
F:271-288/Product: transmembrane #status predicted <TM2>
F:286-778/Product: major envelope protein E #status predicted <MEE>
F:736-753/Product: transmembrane #status predicted <TM3>
F:756-778/Product: transmembrane #status predicted <TM4>
F:779-1187/Product: nonstructural protein NS1 #status predicted <NS1>
F:1133-1151/Product: transmembrane #status predicted <TM5>
F:1160-1179/Product: transmembrane #status predicted <TM6>
F:1188-1154/Product: nonstructural protein NS2a #status predicted <N2a>
F:1355-1484/Product: nonstructural protein NS2b #status predicted <N2b>
F:1485-2107/Product: nonstructural protein NS3 #status predicted <NS3>
F:1682-1689/Region: nucleotide-binding motif A (P-loop)
F:1769-1774/Region: nucleotide-binding motif B
F:1773-1776/Region: DEAH motif
F:2108-2394/Product: nonstructural protein NS4a #status predicted <N4a>
F:2395-2506/Product: nonstructural protein NS4b #status predicted <N4b>
F:2507-3411/Product: nonstructural protein NS5 #status predicted <NS5>
F:134, 150, 172, 594, 908, 986, 1796, 2062, 2320, 2346, 2408, 2467, 2720, 2734, 2740/Binding site: capsid protein

Query Match 52.7% Score 49; DB 1; Length 3411;
Best Local Similarity 50.0%; Pred. NO. 2.8e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LKRFNRKIKRKGOK 18
:| |||||:|:|:|:
Db 20 VRSLSMKIKOKTKOIGNR 37

RESULT 12

772227 conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: C72227
R:Neilson, K.E.; Clayton, R.A.; Gill, S.R.; Gain, M.L.; Dodson, R.J.; Hatf, D.H.; Hickey, C.M.
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
A:Reference number: A72200; MUID:99287316
A:Accession: C72227
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-152 <ARN>
A:Cross-references: GB:AE001807; GB:AE000512; NID:g4982216; PIDN:AAD36713.1; PID:g4982222
A:Experimental source: strain MSB8
A:Genetics:
A:Gene: TMI646

```

Query Match          51.6%; Score 48; DB 2; Length 152;
Best Local Similarity 52.6%; Pred. No. 21;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

      1 LKFNKIKKEKLKKGOKI 19
      |:::| ||| | | | |
      Db 74 LKRYNKAIFKFLKLEKKI 92

RESULT 13
T04951
hypothetical protein F7J7.140 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 26-Aug-1999
C:Accession: T04951
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.
submitted to the Protein Sequence Database, July 1998
A:Reference number: 215391
A:Accession: T04951
A:Molecule type: DNA
A:Residues: 1293 <BEV>
A:Cross-references: EMBL:AL021960
A:Experimental source: cultivar Columbia; BAC clone F7J7
C:Genetics:
A:Map position: 4
A:Introns: 107/2; 214/3
A:Note: F7J7.140
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match          51.6%; Score 48; DB 2; Length 293;
Best Local Similarity 44.4%; Pred. No. 39;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

      2 KFRNKIKKEKLKKGOKI 19
      |::|:::|::| | |
      Db 266 REFROQVEDVKKFGFKV 283

RESULT 14
E90061
hypothetical protein SA2351 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: E90061
R:Kurouda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Cui, A.; Mizutani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, H.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: E90061
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1497 <KOR>
A:Cross-references: GB:BA000018; PID:913702514; PIDN:BAB43655.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA2351

Query Match          51.6%; Score 48; DB 2; Length 497;
Best Local Similarity 64.3%; Pred. No. 64;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

      3 KFRNKIKKEKLKKG 16
      |::| | |::|::|
      Db 393 KFRDKILDKLEKMG 406

RESULT 15
569444

```

conserved hypothetical protein AF1556 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 08-Oct-1999
C:Accession: C69444
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343
A:Accession: C69444
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-274 <KLE>
A:Cross-references: GB:AE000995; GB:AE000782; NID:g2689318; PIDN:AAB89691.1; PID:g264900

Query Match 51.1%; Score 47.5; DB 2; Length 274;
Best Local Similarity 57.9%; Pred. No. 43;
Matches 11; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 1 LRRFRN--KIKELKKIG 16
||| |:::|||||
Db 236 LRKLCAKVRDKLKKIG 254

Search completed: July 12, 2002, 07:56:58
Job time: 225 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2002, 08:04:42 ; Search time 46.15 Seconds
(without alignments)
15.941 Million cell updates/sec

Title: US-09-642-744B-21
Perfect score: 93
Sequence: 1 LRFRNKKIKELKKIGQKI 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	93	100.0	171	1	CP18_RABIT
2	64	68.8	173	1	GRAM_MOUSE
3	52	55.9	1111	1	YJEP_HAIEIN
4	50	53.8	811	1	COAT_CMDNV
5	49	52.7	1163	1	POLG_YEYFV
6	49	52.7	3411	1	POLG_YEYFV
7	49	52.7	3411	1	POLG_YEYFV
8	47	50.5	288	1	ST1A_BOVIN
9	47	50.5	288	1	ST1A_HUMAN
10	47	50.5	288	1	ST1A_MOUSE
11	47	50.5	288	1	ST1A_MOUSE
12	46	49.5	166	1	MP36_PIG
13	46	49.5	169	1	DEF_VIBCH
14	46	49.5	288	1	ST1B_BOVIN
15	46	49.5	288	1	ST1B_HUMAN
16	46	49.5	288	1	ST1B_MOUSE
17	45	48.4	165	1	MB34_BOVIN
18	45	48.4	486	1	RFG3_YEAST
19	45	48.4	809	1	COAT_DSDNV
20	45	48.4	1630	1	MSP1_PLAFLW
21	45	48.4	1639	1	MSP1_PLAFLW
22	44	47.3	109	1	Y639_METJA
23	44	47.3	175	1	Y240_METJA
24	44	47.3	188	1	Y460_PYRHO
25	44	47.3	356	1	UPPS_METJA
26	44	47.3	356	1	UPPS_METJA
27	44	47.3	903	1	DPOL_BPR69
28	44	47.3	2198	1	YLJ2_CAEBL
29	43	46.2	66	1	Y766_METJA
30	43	46.2	96	1	RS20_THEMEA
31	43	46.2	117	1	PFDB_PYRAB
32	43	46.2	117	1	PFDB_PYRHO
33	43	46.2	176	1	YL57_ARCFU

34	43	46.2	188	1	VF59_METJA	058954 methanococc
35	43	46.2	194	1	Y182_METJA	057641 methanococc
36	43	46.2	281	1	BLAC_STRAU	P00807 Staphylococ
37	43	46.2	290	1	STX_APLCA	016932 aplysia cal
38	43	46.2	308	1	Y227_METJA	057680 methanococc
39	43	46.2	346	1	Y21C_MYCGE	092b78 mycoplasma
40	43	46.2	555	1	ILVD_AQUAE	067009 aquifex aeo
41	43	46.2	629	1	DNAC_PAVLU	P30722 pavlova lut
42	43	46.2	634	1	YC36_METJA	058633 methanococc
43	43	46.2	835	1	Y422_MYCGE	P47661 mycoplasma
44	43	46.2	1207	1	DPOL_ASFB7	P42489 african swi
45	43	46.2	1244	1	DPOL_ASFL6	P43139 african swi

ALIGNMENTS

```

RESULT 1
ID CP18_RABIT STANDARD; PRT; 171 AA.
AC P25230;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Antimicrobial protein CAP18 precursor (18 kDa lipopolysaccharide-
  binding protein) (18 kDa cationic protein) (CAP18-A).
GN CAP18.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 135-159.
RC TISSUE=bone marrow; PubMed=1883348;
RX MEDLINE=91354246; PubMed=1883348;
RA Larrick J.W., Morgan J.G., Pallings I., Hirata M., Yen M.H.;
RT "Complementary DNA sequence of rabbit CAP18 -- a unique
  lipopolysaccharide binding protein."
RL Biochem. Biophys. Res. Commun. 179:170-175(1991).
RN [2]
RP SEQUENCE OF 135-159, AND CHARACTERIZATION.
RX MEDLINE=94178952; PubMed=8132348;
RA Hirata M., Shimomura Y., Yoshida M., Morgan J.G., Pallings I.,
  Wilson D., Yen M.H., Wright S.C., Larrick J.W.;
RT "Characterization of a rabbit cationic protein (CAP18) with
  lipopolysaccharide-inhibitory activity."
RL Infect. Immun. 62:1421-1426(1994).
RN [3]
RP SEQUENCE OF 135-154, AND CHARACTERIZATION.
RX MEDLINE=94075827; PubMed=8254193;
RA Larrick J.W., Hirata M., Zheng H., Zhong J., Bolin D.,
  Cavallion J.-M., Warren H.S., Wright S.C.;
RT "A novel granulocyte-derived peptide with lipopolysaccharide-
  neutralizing activity."
RL J. Immunol. 152:231-240(1994).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=94148064; PubMed=8313956;
RA Tossi A., Scoocchi M., Skerlavaj B., Gennaro R.;
RT "Identification and characterization of a primary antibacterial
  domain in CAP18, a lipopolysaccharide binding protein from rabbit
  leukocytes."
RL FEBS Lett. 339:108-112(1994).
RN [5]
RP STRUCTURE BY NMR OF 135-166.
RX MEDLINE=95377455; PubMed=7649303;
RA Chen C., Brock R., Luh F., Chou P.-J., Larrick J.W., Huang R.-F.,
  Huang T.-H.;
RT "The solution structure of the active domain of CAP18 -- a
  lipopolysaccharide binding protein from rabbit leukocytes."
RL FEBS Lett. 370:46-52(1995).
CC -I- FUNCTION: CAP18 BINDS TO THE LIPID A MOIETY OF BACTERIAL
  LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER

```

```

CC      MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. HAS ANTIBIOTIC ACTIVITY.
CC      -1- TISSUE SPECIFICITY: NEUTROPHILS.
CC      -1- SIMILARITY: BELONGS TO THE CATHHELICIDIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: M73998; AAA31187.1; -.
CC      PIR: J01171; J01171.
CC      PDB: 1LTP; 31-MAR-95.
CC      InterPro: IPR001894; Cathelicidin.
CC      Pfam: PF00666; Cathelicidins; 1.
CC      ProDom: PD001838; Cathelicidins; 1.
CC      PROSITE: PS00946; CATHHELICIDINS_1; 1.
CC      PROSITE: PS00947; CATHHELICIDINS_2; 1.
CC      Antibiotic; Signal; 3D-structure.
CC      SIGNAL 1 29
CC      CHAIN 30 171 ANTIMICROBIAL PROTEIN CAP18.
CC      CHAIN 135 171 ANTIMICROBIAL PROTEIN CAP7.
CC      MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID
CC      (BY SIMILARITY).
CC      DISULFID 85 96 BY SIMILARITY.
CC      DISULFID 107 124 BY SIMILARITY.
CC      VARIANT 157 157 K -> D.
CC      SEQUENCE 171 AA; 19805 MW; D7BF2103BCFB13C4 CRC64;

Query Match 100.0%; Score 93; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 2.3e-05;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRRFRNRIKELKIGOKI 19
DB 140 LRRFRNRIKELKIGOKI 158

RESULT 2
GRAM_MOUSE STANDARD; PRT: 173 AA.
AC PS1437.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cathelin-related antimicrobial peptide precursor (cramp) (Cathelin-
DE like protein) (CLP).
GN CNLP OR CRAMP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Bone marrow;
RX MEDLINE=96326596; PubMed=8706928;
RA Popsuena A.E., Zinovjeva M.V., Visser Y.W.M., Fibbe W.E.,
RA Belyavsky A.V.;
RA "A novel murine cathelin-like protein expressed in bone marrow.";
RA FEBS Lett. 391:5-8(1996).
RN [3]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ.
RA Huttner K.M., Pirano J., Gallo R.L.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTS AS A POTENT ANTIMICROBIAL PEPTIDE.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN TESTIS, SPLEEN, STOMACH, AND
CC INTESTINE. VERY LOW EXPRESSION FOUND IN HEART, LUNG AND SKELETAL
CC MUSCLE. NO EXPRESSION IN BRAIN, KIDNEY OR LIVER.
CC -1- SIMILARITY: BELONGS TO THE CATHHELICIDIN FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: U43409; AA85898.1; -.
CC      EMBL: X94353; CA64078.1; -.
CC      EMBL: AF035680; AAB88303.1; -.
CC      HSSP: P25230; 1LTP.
CC      MGD: MGI:108443; Cnlp.
CC      InterPro: IPR001894; Cathelicidin.
CC      Pfam: PF00666; Cathelicidins; 1.
CC      ProDom: PD001838; Cathelicidins; 1.
CC      PROSITE: PS00946; CATHHELICIDINS_1; 1.
CC      PROSITE: PS00947; CATHHELICIDINS_2; 1.
CC      Antibiotic; Signal.
CC      SIGNAL 1 27
CC      PROPEP 28 139 POTENTIAL.
CC      CHAIN 140 173 CATHELIN-RELATED ANTIMICROBIAL PEPTIDE.
CC      MOD_RES 28 28 PYRROLIDONE CARBOXYLIC ACID (BY
CC      SIMILARITY).
CC      DISULFID 83 94 BY SIMILARITY.
CC      DISULFID 105 122 BY SIMILARITY.
CC      CONFLICT 24 24 L -> M (IN REF. 2).
CC      CONFLICT 173 173 MISSING (IN REF. 1).
CC      SEQUENCE 173 AA; 19581 MW; 143F0E784762F77E CRC64;

Query Match 68.8%; Score 64; DB 1; Length 173;
Best Local Similarity 78.9%; Pred. No. 0.12;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LRRFRNRIKELKIGOKI 19
DB 142 LRRGEGIKELKIGOKI 160

RESULT 3
YJEP_HAEIN STANDARD; PRT: 1111 AA.
ID YJEP_HAEIN
AC 057362.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein H10195.1 precursor.
DE H10195.1.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

```

RA Fine L.D., Fritschman J.L., Fuhrmann J.L., Geochagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae Rd.";
 RL Science 269:496-512(1995).
 RP [2]
 RP IDENTIFICATION BY MASS SPECTROMETRY.
 RX MEDLINE-20137488: PubMed10675023:
 RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
 RA Gray C., Fountoulakis M.;
 RT "Two-dimensional map of the proteome of Haemophilus influenzae";
 RL Electrophoresis 21:411-429(2000).
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -I- SIMILARITY: BELONGS TO THE UPF003 FAMILY. STRONG, TO E.COLI YJEP.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32704; AAC21864.1; -;
 DR TIGR: H10195.1; -;
 DR InterPro: IPR001880: UPF0003.
 DR Pfam: PF00924; MS-channel; 1.
 DR PROSITE: PS01246; UPF0003; 1.
 KW Transmembrane; Signal; Complete proteome.
 FT SIGNAL 1 31
 FT CHAIN 32 1111 POTENTIAL.
 FT TRANSMEM 490 510 POTENTIAL.
 FT TRANSMEM 538 558 POTENTIAL.
 FT TRANSMEM 572 592 POTENTIAL.
 FT TRANSMEM 620 640 POTENTIAL.
 FT TRANSMEM 644 664 POTENTIAL.
 FT TRANSMEM 694 714 POTENTIAL.
 FT TRANSMEM 797 817 POTENTIAL.
 FT TRANSMEM 840 860 POTENTIAL.
 FT TRANSMEM 885 905 POTENTIAL.
 FT TRANSMEM 922 942 POTENTIAL.
 FT TRANSMEM 1003 1023 POTENTIAL.
 FT SEQUENCE 1111 AA; 126825 MW; 15DF8F9F7E4E701C CRC64;
 SQ
 Query Match 55.9%; Score 52; DB 1; Length 1111;
 Best Local Similarity 52.9%; Pred. No. 19;
 Matches 9; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
 Oy 3 KFRNKIKKKIKGKI 19
 Db 509 KFRNKIKQDLNKRIRI 525
 RESULT 4
 COAT_GMDNV STANDARD; PRT; 811 AA.
 ID Q90125; Q90128; Q90126; Q90127;
 AC Q90125; Q90128; Q90126; Q90127;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Coat protein VP1 (Structural protein VP1) [contains: Coat protein VP
 DE (Structural protein VP2); Coat protein VP3 (Structural protein VP3);
 DE Coat protein VP4 (Structural protein VP4)].
 GN VP.
 OS Galliera mellonella densovirus (GMDNV).
 OC Viruses; ssDNA viruses; Parvoviridae; Densovirinae; Densovirus.
 OX NCBI_TaxID=37138;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Tjissen P.;
 RT "Organization and expression of the ambisense genome of

RT densoonucleosis virus of Galliera mellonella (GMDNV).";
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (3.6 ANGSTROMS) OF 396-811.
 RX MEDLINE-99036856; PubMed-9817847;
 RA Simpson A.A., Chipman P.R., Baker T.S., Tjissen P., Rossmann M.G.;
 RT "The structure of an insect parvovirus (Galliera mellonella
 RT densovirus) at 3.7 A resolution.";
 RL Structure 6:1355-1367(1998).
 CC -I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
 CC OF A COMBINATION OF VP4, VP3, VP2 AND VP1.
 CC -I- ALTERNATIVE PRODUCTS: THE FOUR DIFFERENT COAT PROTEINS ARE
 CC PRODUCED BY ALTERNATIVE INITIATION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L32896; AAA66966.1; -;
 DR EMBL: L32896; AAA66964.1; -;
 DR EMBL: L32896; AAA66965.1; -;
 DR EMBL: L32896; AAA66967.1; -;
 DR PDB: 1DNV; 16-FEB-99.
 DR InterPro: IPR003433; denso_VP4.
 DR Pfam: PF02336; denso_VP4; 1.
 KW Coat protein; Alternative initiation; 3D-structure.
 FT CHAIN 1 811
 FT CHAIN 277 811 COAT PROTEIN VP1.
 FT CHAIN 324 811 COAT PROTEIN VP2.
 FT CHAIN 375 811 COAT PROTEIN VP3.
 FT CHAIN 811 811 COAT PROTEIN VP4.
 FT SEQUENCE 811 AA; 88542 MW; 0A61B09172DF99B5 CRC64;
 SQ
 Query Match 53.8%; Score 50; DB 1; Length 811;
 Best Local Similarity 58.8%; Pred. No. 26;
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 3 KFRNKIKKKIKGKI 19
 Db 97 KFRNKIKKKIKGKI 113
 RESULT 5
 POLG_YEFV8 STANDARD; PRT; 1163 AA.
 ID P29165;
 AC P29165;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polyprotein [contains: Capsid protein C (Core protein); Matrix
 DE protein (Envelope protein M); Major envelope protein E; Nonstructural
 DE proteins NS1, NS2A] (Fragment).
 OS Yellow fever virus (strain 1899/81).
 OC Viruses; ssRNA positive-strand viruses; no DNA stage; Flaviviridae;
 OC Flavivirus.
 OX NCBI_TaxID=31641;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91011358; PubMed-2145394;
 RA Ballinger-Crabtree M.E., Miller B.R.;
 RT "Partial nucleotide sequence of South American yellow fever virus
 RT strain 1899/81: structural proteins and NS1.";
 RL J. Gen. Virol. 71:2115-2121(1990).
 CC -I- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D14458; BAA03355.1; -
DR PIR: J00374; GNMVY8.
DR HSP: P14336; ISVB.
DR InterPro: IPR000069; Flavi_M.
DR InterPro: IPR001157; Flavi_NS1.
DR InterPro: IPR001122; Flavi_capsid.
DR InterPro: IPR000336; Flavi_glycoprote.
DR InterPro: IPR002535; Flavi_propep.
DR Pfam: PF01003; Flavi_capsid; 1.
DR Pfam: PF00869; Flavi_glycoprot; 1.
DR Pfam: PF02832; Flavi_glycoprot; 1.
DR Pfam: PF01004; Flavi_M; 1.
DR Pfam: PF00948; Flavi_NS1; 1.
DR Pfam: PF01570; Flavi_propep; 1.
DR ProDom: PD001496; Flavi_NS1; 1.
DR ProDom: PD001556; Flavi_glycoprote; 1.
KM Polyprotein: Glycoprotein: Core protein; Coat protein;
FT Envelope protein; Transmembrane; Nonstructural protein.
FT INIT_MET 1 1

FT CHAIN 1 121
FT PROPEP 122 210
FT CHAIN 211 285
FT CHAIN 286 778
FT CHAIN 779 1130
FT CHAIN 1131 >1163
FT TRANSMEM 249 269
FT TRANSMEM 271 285
FT TRANSMEM 740 753
FT TRANSMEM 755 778
FT TRANSMEM 1159 >1163
FT DISULFID 288 315
FT DISULFID 345 401
FT DISULFID 359 390
FT DISULFID 377 406
FT DISULFID 467 568
FT DISULFID 585 615
FT CARBOHYD 114 134
FT CARBOHYD 150 150
FT CARBOHYD 554 554
FT CARBOHYD 908 908
FT CARBOHYD 986 986
FT NON_TER 1163 1163
SQ SEQUENCE 1163 AA; 128520 MW; 6BEBDA317D722E01 CRC64;

Query Match 52.7%; Score 49; DB 1; Length 1163;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKRNKIKKKIKG 18
: 1 1111111111:
DB 20 VRSLNKKIKOKTKQIGNR 37

RESULT 6
POLG_YEYV1 STANDARD: PRT; 3411 AA.
ID POLG_YEYV1 042028:
AC 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [contains: Capsid protein C (Core protein); Matrix
DE protein (Envelope protein M); Major envelope protein E; Nonstructural
DE proteins NS1, NS2A, NS2B, NS4A and NS4B; Protease/helicase
DE (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)]

DE (NS5)}.
OS Yellow fever virus (strain 17D).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=11090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85272570; Pubmed=4023707;
RA Rice C.M., Ienches E.M., Eddy S.R., Shin S.J., Sheets R.L.,
RA Strauss J.H.;
RT "Nucleotide sequence of yellow fever virus: implications for
RT flavivirus gene expression and evolution";
RL Science 229:726-733(1985).
CC - FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC - CATALYTIC ACTIVITY: HYDROLYSIS OF FOUR PEPTIDE BONDS IN THE VIRAL
CC PRECURSOR POLYPROTEIN, COMMONLY WITH ASP OR GLU IN THE P6
CC POSITION, CYS OR THR IN P1 AND SER OR ALA IN P1'.
CC - SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X03700; CA27332.1; -
DR PIR: A03914; GNMVY.
DR HSP: P14336; ISVB.
DR MEROPS: S07.001; -
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR000069; Flavi_M.
DR InterPro: IPR001157; Flavi_NS1.
DR InterPro: IPR000752; Flavi_NS2A.
DR InterPro: IPR000487; Flavi_NS2B.
DR InterPro: IPR000404; Flavi_NS4A.
DR InterPro: IPR001528; Flavi_NS4B.
DR InterPro: IPR000208; Flavi_NS5.
DR InterPro: IPR001122; Flavi_capsid.
DR InterPro: IPR000336; Flavi_glycoprote.
DR InterPro: IPR001850; Flavi_helicase.
DR InterPro: IPR002535; Flavi_propep.
DR InterPro: IPR002877; FtsJ.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF01003; Flavi_capsid; 1.
DR Pfam: PF00869; Flavi_glycoprot; 1.
DR Pfam: PF02832; Flavi_glycoprot; 1.
DR Pfam: PF00949; Flavi_helicase; 1.
DR Pfam: PF01004; Flavi_M; 1.
DR Pfam: PF00948; Flavi_NS1; 1.
DR Pfam: PF01005; Flavi_NS2A; 1.
DR Pfam: PF01002; Flavi_NS2B; 1.
DR Pfam: PF01350; Flavi_NS4A; 1.
DR Pfam: PF01349; Flavi_NS4B; 1.
DR Pfam: PF00972; Flavi_NS5; 1.
DR Pfam: PF01570; Flavi_propep; 1.
DR Pfam: PF01728; FtsJ; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR ProDom: PD001496; Flavi_NS1; 1.
DR ProDom: PD001556; Flavi_glycoprote; 1.
DR SMART: SM00490; HELIC_C; 1.
KM Polyprotein: Glycoprotein; Transferrase; RNA-directed RNA polymerase;
KM core protein; coat protein; Envelope protein; Hydrolase; Helicase;
KM ATP-binding; Transmembrane; Nonstructural protein.
FT INIT_MET 1 1
FT CHAIN 1 121
FT PROPEP 122 210
FT CHAIN 211 285
FT CHAIN 286 778
FT CHAIN 779 1130
FT CHAIN 1131 >1163
FT TRANSMEM 249 269
FT TRANSMEM 271 285
FT TRANSMEM 740 753
FT TRANSMEM 755 778
FT TRANSMEM 1159 >1163
FT DISULFID 288 315
FT DISULFID 345 401
FT DISULFID 359 390
FT DISULFID 377 406
FT DISULFID 467 568
FT DISULFID 585 615
FT CARBOHYD 114 134
FT CARBOHYD 150 150
FT CARBOHYD 554 554
FT CARBOHYD 908 908
FT CARBOHYD 986 986
FT NON_TER 1163 1163
SQ SEQUENCE 1163 AA; 128520 MW; 6BEBDA317D722E01 CRC64;

Query Match 52.7%; Score 49; DB 1; Length 1163;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 LKRNKIKKKIKG 18
: 1 1111111111:
DB 20 VRSLNKKIKOKTKQIGNR 37

RESULT 6
POLG_YEYV1 STANDARD: PRT; 3411 AA.
ID POLG_YEYV1 042028:
AC 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [contains: Capsid protein C (Core protein); Matrix
DE protein (Envelope protein M); Major envelope protein E; Nonstructural
DE proteins NS1, NS2A, NS2B, NS4A and NS4B; Protease/helicase
DE (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)]

```

FT PROPEP      122      210
FT CHAIN        211      285
FT CHAIN        286      778
FT CHAIN        779      1130
FT CHAIN        1131     1354
FT CHAIN        1355     1484
FT CHAIN        1485     2107
FT CHAIN        2108     2256
FT CHAIN        2257     2506
FT CHAIN        2507     3411
FT TRANSMEM     249      269
FT TRANSMEM     271      285
FT TRANSMEM     740      753
FT TRANSMEM     755      778
FT TRANSMEM     1159     1180
FT DOMAIN       383      396
FT NP_BIND     1682     1689
FT SITE        1773     1776
FT DISULFID     288      315
FT DISULFID     345      401
FT DISULFID     359      390
FT DISULFID     377      406
FT DISULFID     467      568
FT DISULFID     585      615
FT CARBOHYD     134      134
FT CARBOHYD     150      150
FT CARBOHYD     908      908
FT CARBOHYD     986      986
FT CARBOHYD     2320     2320
FT CARBOHYD     2346     2346
FT CARBOHYD     2467     2467
SQ SEQUENCE     3411 AA; 379512 MW; 680E0FACD23DCFA6 CRC64;

```

```

Query Match      52.7%; Score 49; DB 1; Length 3411;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

OY      1 LRKFRNKIKKIKTGOK 18
      :| ||||| :|||:
Db      20 VRSLNSKIKOKTKOIGNR 37

```

```

RESULT      7
ID POLG_YEFV2 STANDARD; PRT: 3411 AA.
AC P19901;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypotein [Contains: Capsid protein C (Core protein); Matrix
DE protein (Envelope protein M); Major envelope protein E; Nonstructural
DE proteins NS1, NS2A, NS2B, NS4A and NS4B; Protease/helicase
DE (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)
DE (NS5)].
OS Yellow fever virus (strain Pasteur 17D-204).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=11091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89282413; PubMed=2734112;
RA Dupuy A., Despres P., Cahour A., Girard M., Bouloy M.;
RT "Nucleotide sequence comparison of the genome of two 17D-204 yellow
RT fever vaccines.";
RL Nucleic Acids Res. 17:3989-3989(1989).
CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF FOUR PEPTIDE BONDS IN THE VIRAL
CC PRECURSOR POLYPROTEIN, COMMONLY WITH ASP OR GLU IN THE P6
CC POSITION, CYS OR THR IN P1 AND SER OR ALA IN P1'.
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A

```

```

CC CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC CC PROTEIN C AND MRNA.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL; X15062; CAB37419.1;
DR PIR; S07757; GMVVP.
DR HSSP; P14336; LSVB.
DR MEROPS; S07.001; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000069; Flav1_M.
DR InterPro; IPR001157; Flav1_NS1.
DR InterPro; IPR000752; Flav1_NS2A.
DR InterPro; IPR000487; Flav1_NS2B.
DR InterPro; IPR000404; Flav1_NS4A.
DR InterPro; IPR001528; Flav1_NS4B.
DR InterPro; IPR000208; Flav1_NS5.
DR InterPro; IPR001122; Flav1_capsid.
DR InterPro; IPR000336; Flav1_glycoprote.
DR InterPro; IPR001850; Flav1_helicase.
DR InterPro; IPR002535; Flav1_Prophep.
DR InterPro; IPR002877; FtsJ.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00869; Flav1_glycoprot; 1.
DR Pfam; PF02832; Flav1_glycop_C; 1.
DR Pfam; PF00949; Flav1_helicase; 1.
DR Pfam; PF01004; Flav1_M; 1.
DR Pfam; PF00948; Flav1_NS1; 1.
DR Pfam; PF01005; Flav1_NS2A; 1.
DR Pfam; PF01002; Flav1_NS2B; 1.
DR Pfam; PF01350; Flav1_NS4A; 1.
DR Pfam; PF01349; Flav1_NS4B; 1.
DR Pfam; PF00972; Flav1_NS5; 1.
DR Pfam; PF01570; Flav1_Prophep; 1.
DR Pfam; PF01728; FtsJ; 1.
DR Pfam; PF00271; helicase_C; 1.
DR ProDom; PD001496; Flav1_NS1; 1.
DR ProDom; PD001556; Flav1_glycoprote; 1.
DR SMART; SM00490; Helic; 1.
DR Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Hydrolyase; Helicase;
KW ATP-binding; Transmembrane; Nonstructural protein.
KW INIT_MET 1
FT CHAIN 1 121
FT PROPEP 122 210
FT CHAIN 211 285
FT CHAIN 286 778
FT CHAIN 779 1130
FT CHAIN 1131 1354
FT CHAIN 1355 1484
FT CHAIN 1485 2107
FT CHAIN 2108 2256
FT CHAIN 2257 2506
FT CHAIN 2507 3411
FT CHAIN 3411 3411
FT NP_BIND 1682 1689
FT SITE 1773 1776
FT TRANSMEM 249 269
FT TRANSMEM 271 285
FT TRANSMEM 740 753
FT TRANSMEM 755 778
FT TRANSMEM 1159 1180
FT DISULFID 288 315
FT DISULFID 345 401
FT DISULFID 359 390

```



```
FT DISULFID 377 406 BY SIMILARITY.
FT DISULFID 467 568 BY SIMILARITY.
FT DISULFID 585 615 BY SIMILARITY.
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 150 150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 986 986 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2320 2320 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2346 2346 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2467 2467 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 3411 AA; 379525 MM; 3298C0771FED23F7 CRC64;
```

```
Query Match 52.7%; Score 49; DB 1; Length 3411;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
OY 1 LRKFRNKIKKIKKIGOK 18
:| |||||:|:|:|
Db 20 VRSLSKIKIKOKTKOICNR 37
```

```
RESULT 8
STLA_BOVIN STANDARD: PRT; 288 AA.
AC P32850;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Syntaxin 1A (Synaptotagmin associated 35 kDa protein) (P35A)
DE (Neuron-specific antigen HPC-1).
GN STX1A.
```

```
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE=92412081; PubMed=1530610;
RA Inoue A., Akagawa K.;
RT "SNP receptors implicated in vesicle targeting and fusion.";
RT "Neuron-specific antigen HPC-1 from bovine brain reveals strong
RT morphology to epimorphin, an essential factor involved in epithelial
RT morphogenesis: identification of a novel protein family.";
RN Biochem. Biophys. Res. Commun. 187:1144-1150(1992).
```

```
RN [2]
RP SEQUENCE OF 29-40; 43-55; 97-106; 160-166 AND 213-224.
RX MEDLINE=93205116; PubMed=8455717;
```

```
RA Soellner T., Whiteheart S.W., Brunner M., Erdjument-Bromage H.,
RA Geramios S., Tempst P., Rothman J.E.;
RT "SNP receptors implicated in vesicle targeting and fusion.";
RT Nature 362:318-324(1993).
```

```
CC -1- FUNCTION: POTENTIALLY INVOLVED IN DOCKING OF SYNAPTIC VESICLES AT
CC PRESYNAPTIC ACTIVE ZONES. MAY PLAY A CRITICAL ROLE IN
CC NEUROTRANSMITTER EXOCYTOSIS.
```

```
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.
```

```
DR HSSP; P32851; J01615.
DR InterPro; IPR000077; T-SNARE.
DR InterPro; IPR000727; T-SNARE.
DR Pfam; PF00804; Syntaxin; 1.
DR SMART; SM00397; T-SNARE; 1.
DR PROSITE; PS00914; SYNTAXIN; 1.
KW Neurotransmitter transport; Coiled coil; Transmembrane; Antigen.
FT DOMAIN 68 109 COILED COIL (POTENTIAL).
FT TRANSMEM 266 288 POTENTIAL.
SQ SEQUENCE 288 AA; 33091 MM; 97ED781DEDEBBA09 CRC64;
```

```
Query Match 50.5%; Score 47; DB 1; Length 288;
```

```
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
```

```
OY 1 LRKFRNKIKKIKKIGOKI 19
:| |||:| | | |
Db 82 IKTKANKVRSLKSIKSI 100
```

```
RESULT 9
STLA_HUMAN STANDARD: PRT; 288 AA.
AC 016623; Q12936; O15447; O15448; Q9BP26;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Syntaxin 1A (Neuron-specific antigen HPC-1).
GN STX1A.
```

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-Brain;
RX MEDLINE=95347620; PubMed=7622072;
RA Zhang R.-D., Maksymowych A.B., Simpson L.L.;
RT "Cloning and sequence analysis of a cDNA encoding human syntaxin 1A,
RT a polypeptide essential for exocytosis.";
RN Gene 159:293-294(1995).
```

```
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
```

```
RC TISSUE-Brain;
RA Fujiwara T., Genda M., Akagawa K.;
RN Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
```

```
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=97456437; PubMed=9311751;
RA Osborne L.R., Soder S., Shi X.-M., Pober B., Costa T., Scherer S.W.,
RA Tsui L.-C.;
RT "Hemizygous deletion of the syntaxin 1A gene in individuals with
RT Williams syndrome.";
RN Am. J. Hum. Genet. 61:449-452(1997).
```

```
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND REVISION TO 144.
```

```
RA Wu Y.-Q., Osborne L.R., Tsui L.-C., Shaffer L.G.;
RT "Refinement of the genomic structure of STX1A and mutation analysis in
RT nondeletion Williams syndrome patients.";
RN Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
```

```
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
```

```
RC TISSUE-Fat;
RX MEDLINE=97157033; PubMed=9003414;
RA Jagadeish M.N., Tellam J.T., Macaulay S.L., Gough K.H., James D.E.,
RA Ward C.W.;
RT "Novel isoform of syntaxin 1 is expressed in mammalian cells.";
RN Biochem. J. 321:151-156(1997).
```

```
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 3).
```

```
RC TISSUE-Lung;
RA Strausberg R.;
RN Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
```

```
RN [7]
RP SEQUENCE OF 15-288 FROM N.A. (ISOFORM 1).
```

```
RC TISSUE-Brain;
RA Zhang R.-D., Simpson L.;
RN Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
```

```
RN [8]
RP TISSUE SPECIFICITY.
RA Botta A., Calza L., Giardino L., Potenza S., Novelli G.,
RA Dallapiccola B.;
RT "Expression analysis of the human HPC-1/syntaxin 1A, a gene deleted
RT in Williams syndrome.";
RN Eur. J. Hum. Genet. Suppl. 7:96-96(1999).
```

```
CC -1- FUNCTION: POTENTIALLY INVOLVED IN DOCKING OF SYNAPTIC VESICLES AT PRESYNAPTIC ACTIVE ZONES. MAY PLAY A CRITICAL ROLE IN NEUROTRANSMITTER EXOCYTOSIS.
CC CC
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Isoform 2 may be soluble.
CC CC
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2/1C and 3; are produced by alternative splicing.
CC CC
CC -1- TISSUE SPECIFICITY: Isoform 1 is highly expressed in embryonic spinal chord and ganglia and in adult cerebellum and cerebral cortex. Isoform 2 is expressed in heart, liver, fat, skeletal muscle, kidney and brain.
CC CC
CC -1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIPHORPHIN FAMILY.
CC CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/cc or send an email to license@sib-sib.ch).
CC CC
DR EMBL: L37792; AAA53519.1; -.
DR EMBL: D37932; BAA07151.1; -.
DR EMBL: U87315; AAK54507.2; -.
DR EMBL: AF297001; AAK54507.2; JOINED.
DR EMBL: AF297002; AAK54507.2; JOINED.
DR EMBL: U87310; AAK54507.2; JOINED.
DR EMBL: AF297003; AAK54507.2; JOINED.
DR EMBL: U87314; AAK54507.2; JOINED.
DR EMBL: U87315; AAB65500.2; -.
DR EMBL: AF297001; AAB65500.2; JOINED.
DR EMBL: AF297002; AAB65500.2; JOINED.
DR EMBL: U87310; AAB65500.2; JOINED.
DR EMBL: AF297003; AAB65500.2; JOINED.
DR EMBL: U87314; AAB65500.2; JOINED.
DR EMBL: BC000444; AAH00444.1; -.
DR EMBL: BC003011; AAH03011.1; -.
DR EMBL: U12918; AAA20940.1; ALT_INIT.
DR HSSP: P32851; IBRO.
DR MIM: 186590; -.
DR InterPro: IPR000017; Syntaxin.
DR InterPro: IPR000727; T-SNARE.
DR Pfam: PF00804; Syntaxin; 1.
DR SMART: SMO0503; Synn; 1.
DR SMART: SMO0397; t-SNARE; 1.
DR PROSITE: PS00914; SYNTAXIN; 1.
KW Neurotransmitter transport; Coiled coil; Transmembrane; Antigen; Alternative splicing.
KW FT DOMAIN 1 265 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 266 286 POTENTIAL.
FT DOMAIN 287 288 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 68 109 COILED COIL (POTENTIAL).
FT VASPLIC 227 288 GEMIDRIEYNEHAVDYERAVSDPTKRAVKYSARRKKM EGMLDREYNEHAVDYERAVSDPTKRAVKYSARRKKM EGMIDREYNEHAVDYERAVSDPTKRAVKYSARRKKM IITCVILGIVASTVGIFA -> TMKRGCLTPRPST IITCVILGIVASTVGIFA E -> V (IN REF. 7).
FT VARSPIC 227 288 RARRAGRS (IN ISOFORM 3).
FT CONFLICT 73 73 E -> V (IN REF. 7).
FT CONFLICT 140 140 D -> V (IN REF. 7).
SO SEQUENCE 288 AA; 33023 MM; 8AC787EEFE65ACA1 CMC64;

Query Match 50.5%; Score 47; DB 1; Length 288;
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
```

ID	ST1A.MOUSE	STANDARD;	PT:	288 AA.
AC	O35526;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Syntaxin 1A (Neuron-specific antigen HPC-1).			
GN	STX1A.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain.			
RA	Fujiwara T., Genda M., Nagai A., Okazaki M., Watanabe T.,			
RA	Nagamatsu S., Akagawa K.;			
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: POTENTIALLY INVOLVED IN DOCKING OF SYNAPTIC VESICLES AT			
CC	PRESYNAPTIC ACTIVE ZONES. MAY PLAY A CRITICAL ROLE IN			
CC	NEUROTRANSMITTER EXOCYTOSIS.			
CC	-1- SUBCELLULAR LOCATION: Membrane-bound.			
CC	-1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.jsb-sib.ch/announce/			
CC	or send an email to license@jsb-sib.ch).			
CC	-----			
DR	EMBL; DA5208; BAA28865.2; -			
DR	HSSP; P32851; IBRO.			
DR	MGI; MGI:109355; Stx1a.			
DR	InterPro; IPR000017; Syntaxin.			
DR	InterPro; IPR000727; T_SNARE.			
DR	Pfam; PF00804; Syntaxin; 1.			
DR	SMART; SM00503; Syntaxin; 1.			
DR	SMART; SM00397; T_SNARE; 1.			
DR	PROSITE; PS00914; SYNTAXIN; 1.			
KW	Neurotransmitter transport; Coiled coil; Transmembrane; Antigen.			
FT	DOMAIN 68 109			
FT	TRANSEM 266 286			
SO	SEQUENCE 288 AA; 33054 MW; 8f8255f2eef86ca CRC64;			
	Query Match	50.5%;	Score 47;	DB 1; Length 288;
	Best Local Similarity	47.4%;	Pred. No. 26;	
	Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;			
QY	1 LKRFNKKIKELKRIKGI 19			
	::: ::			
DB	82 IKKTANKVKRSKLKSIQSI 100			
RESULT 11				
ID	ST1A.RAT	STANDARD;	PT:	288 AA.
AC	P32851;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Syntaxin 1A (Synaptobrevin associated 35 kDa protein) (P35A)			
DE	(Neuron-specific antigen HPC-1).			
GN	STX1A OR SAP.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 8-37 AND 44-102.			
RC	TISSUE=Brain;			
CC	MEDLINE=92268107; PubMed=1587842;			

RA Inoue A., Obata K., Akagawa K.;
 RT "Cloning and sequence analysis of cDNA for a neuronal cell membrane
 RT antigen, HPC-1.";
 RL J. Biol. Chem. 267:10613-10619(1992).
 RN [2]
 RP SEQUENCE OF 4-288 FROM N.A., AND SEQUENCE OF 1-15 AND 95-113.
 RC TISSUE-Brain;
 RX MEDLINE-93094187; PubMed-1334074;
 RA Yoshida A., Oho C., Omori A., Kuvahara R., Ito T., Takahashi M.;
 RT "HPC-1 is associated with synaptotagmin and omega-conotoxin
 RT receptor.";
 RL J. Biol. Chem. 267:24925-24928(1992).
 RN [3]
 RP SEQUENCE OF 4-288 FROM N.A., AND SEQUENCE OF 4-16 AND 31-56.
 RC TISSUE-Brain;
 RX MEDLINE-92335866; PubMed-1321498;
 RA Bennett M.K., Calakos N., Scheller R.H.;
 RT "Syntaxin: a synaptic protein implicated in docking of synaptic
 RT vesicles at presynaptic active zones.";
 RL Science 257:255-259(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93386759; PubMed-7690687;
 RA Bennett M.K., Garcia-Arreaza J.E., Elferink L.A., Peterson K.,
 RA Fleming A.M., Hazuka C.D., Scheller R.H.;
 RT "The syntaxin family of vesicular transport receptors.";
 RL Cell 74:863-873(1993).
 RN [5]
 RP STRUCTURE BY NMR OF 27-146.
 RX MEDLINE-98424252; PubMed-9753330;
 RA Fernandez I., Ubach J., Dulhova I., Zhang X., Suedhof T.C., Rizo J.;
 RT "Three-dimensional structure of an evolutionarily conserved N-terminal
 RT domain of syntaxin 1A.";
 RL Cell 94:841-849(1998).
 CC -1- FUNCTION: POTENTIALLY INVOLVED IN DOCKING OF SYNAPTIC VESICLES AT
 CC PRESYNAPTIC ACTIVE ZONES. MAY PLAY A CRITICAL ROLE IN
 CC NEUROTRANSMITTER EXOCYTOSIS.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound
 CC TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN CEREBRAL CORTEX,
 CC HIPPOCAMPUS, CEREBELLUM, ADRENAL MEDULLA AND RETINA WITH WEAK
 CC EXPRESSION DETECTED IN NON-NEURONAL TISSUES.
 CC -1- PTM: PHOSPHORYLATED BY CK2.
 CC -1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D10392; BAA01231.1; ALT_INIT.
 DR EMBL: D12519; BAA02089.1; -
 DR EMBL: M95734; AAA42195.1; -
 DR PIR: A48213; A48213.
 DR PDB: 1HR0; 02-SEP-98.
 DR InterPro: IPR000017; Syntaxin.
 DR InterPro: IPR000727; T_SNARE.
 DR Pfam: PF00804; Syntaxin.1.
 DR SMART: SM00503; Sytn.1.
 DR SMART: SM00397; T_SNARE.1.
 DR PROSITE: PS00914; SYNTAXIN.1.
 KW Neurotransmitter transport; Coiled coil; Transmembrane; Antigen;
 KW Phosphorylation; 3D-structure.
 FT DOMAIN 13 19 ASP-RICH (ACIDIC).
 FT DOMAIN 68 109 COILED COIL (POTENTIAL).
 FT TRANSMEM 266 288 POTENTIAL.
 SQ SEQUENCE 288 AA; 33067 MW; 8F929F489E323CAA CRC64;

Query Match 50.5%; Score 47; DB 1; Length 288;
 Best Local Similarity 47.4%; Pred. No. 26;

Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 1 LKFRNKKIKELKKIGOKI 19
 DB 82 IKTKANKVRSKLSIKSIEOSI 100
 RESULT 12
 MP36_PIG STANDARD; PRT; 166 AA.
 ID MP36_PIG
 AC P49931;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Antibacterial peptide PMAP-36 precursor (Myeloid antibacterial peptide
 DE 36).
 GN PMAP36.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A., AND SYNTHESIS OF 130-166.
 RC TISSUE=bone marrow;
 RX MEDLINE-94123775; PubMed-8293820;
 RA Storici P., Scocchi M., Tossi A., Gennaro R., Zanetti M.;
 RT "Chemical synthesis and biological activity of a novel antibacterial
 RT peptide deduced from a pig myeloid cDNA.";
 RL FEBS Lett. 337:303-307(1994).
 CC -1- FUNCTION: EXERTS ANTIMICROBIAL ACTIVITY AGAINST BOTH GRAM-POSITIVE
 CC AND NEGATIVE BACTERIA. ITS ACTIVITY APPEARS TO BE MEDIATED BY ITS
 CC ABILITY TO DAMAGE BACTERIAL MEMBRANES.
 CC -1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L29125; AAA31070.1; -
 DR InterPro: IPR001894; Cathelicidin.
 DR Pfam: PF00666; Cathelicidins.1.
 DR Prodom: PD001838; Cathelicidin.1.
 DR PROSITE: PS00946; CATHELICIDINS_1; 1.
 DR PROSITE: PS00947; CATHELICIDINS_2; 1.
 KW Antibiotic; Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT PROPEP 30 129 ANTIBACTERIAL PEPTIDE PMAP-36.
 FT CHAIN 130 166 PYRROLIDONE CARBOXYLIC ACID (BY
 FT MOD_RES 30 30 SIMILARITY).
 FT DISULFID 85 96 BY SIMILARITY.
 FT DISULFID 107 124 BY SIMILARITY.
 SQ SEQUENCE 166 AA; 18647 MW; 94B13C69709DA64B CRC64;

Query Match 49.5%; Score 46; DB 1; Length 166;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 2 RKFRNKKIKELKKIGQ 17
 DB 134 RRLKKTRKRLKKIKG 149
 RESULT 13
 DEF_VIBCH STANDARD; PRT; 169 AA.
 ID DEF_VIBCH
 AC Q9KVU3;
 DT 01-MAR-2002 (Rel. 41, Created)

```

DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Peptide deformylase (EC 3.5.1.88) (PDF) (Polypeptide deformylase).
GN DEF OR VC00046.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=204066833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Ginn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unyam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Melakalos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- FUNCTION: Removes the formyl group from the N-terminal Met of
CC newly synthesized proteins. Requires at least a dipeptide for an
CC efficient rate of reaction. N-terminal L-methionine is a
CC prerequisite for activity but the enzyme has broad specificity at
CC other positions (By similarity).
CC -1- CATALYTIC ACTIVITY: Formyl-L-methionyl peptide + H(2)O = formate +
CC methionyl peptide.
CC -1- COFACTOR: Binds 1 iron(II) ion (By similarity).
CC -1- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004095; AAF93224.1; -
DR TIGR: VC00046; -
DR InterPro: IPR000181; Pep_deformylase.
DR Pfam: PF01327; Pep_deformylase; 1.
DR ProDom: PD003844; Pep_deformylase; 1.
KW Protein biosynthesis; Hydrolase; Iron; Complete proteome.
FT METAL 91 91 IRON (BY SIMILARITY).
FT METAL 133 133 IRON (BY SIMILARITY).
FT ACT_SITE 134 134 BY SIMILARITY.
FT METAL 137 137 IRON (BY SIMILARITY).
SQ SEQUENCE 169 AA; 19147 MW; 16DB00B08CA40FC7 CRC64;

Query Match 49.5%; Score 46; DB 1; Length 169;
Best Local Similarity 81.8%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913, 9940;
RN (1)
RP SEQUENCE FROM N.A.
RC SPECIES=Bovine; TISSUE=Brain;
RA Morita T., Mori H., Sakimura K., Mishina M., Sekine Y.,
RA Tsugita A., Odani S., Horikawa H.P., Saisu H., Abe T.;
RT "Synaptocalinin I, a protein associated with brain omega-conotoxin-
RT sensitive calcium channels."
RL Biomed. Res. 13:357-364(1992).
RN (2)
RP SEQUENCE OF 46-81; 125-137 AND 158-181.
RC SPECIES=Bovine;
RX MEDLINE=93205116; PubMed=8455717;
RA Soellner T., Whiteheart S.W., Brunner M., Erdjument-Bromage H.,
RA Geramans S., Tempst P., Rothman J.E.;
RT "SNAP receptors implicated in vesicle targeting and fusion."
RL Nature 362:318-324(1993).
RN (3)
RP SEQUENCE FROM N.A.
RC SPECIES=Sheep; TISSUE=Brain;
RA Helps C.R., Harbour D.A.;
RT "Cloning and sequence analysis of sheep syntaxin 1B."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POTENTIALLY INVOLVED IN DOCKING OF SYNAPTIC VESICLES
CC AT PRESYNAPTIC ACTIVE ZONES.
CC -1- SUBCELLULAR LOCATION: Membrane-bound (potential).
CC -1- SIMILARITY: BELONGS TO THE SYNTAXIN/EPIMORPHIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14133; BAA03188.1; -
DR EMBL: AF081781; AAC31961.1; -
DR HSSP: P32851; 1BR0.
DR InterPro: IPR000017; Syntaxin.
DR InterPro: IPR000727; T-SNARE.
DR Pfam: PF00804; Syntaxin; 1.
DR SMART: SM00503; Synn; 1.
DR SMART: SM00397; L-SNARE; 1.
DR PROSITE: PS00914; SYNTAXIN; 1.
KW Neurotransmitter transport; Coiled coil; Transmembrane.
FT DOMAIN 29 104 COILED COIL (POTENTIAL).
FT TRANSMEM 265 288 POTENTIAL.
FT CONFLICT 79 79 T -> A (IN REF. 2).
SQ SEQUENCE 288 AA; 33274 MW; C66D4785F63C81F CRC64;

Query Match 49.5%; Score 46; DB 1; Length 288;
Best Local Similarity 47.4%; Pred. No. 35;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Fujiwara T., Genda M., Akagawa K.;
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases
CC -! FUNCTION: POTENTIALLY INVOLVED IN DOCKING OF SYNAPTIC VESICLES
CC AT PRESYNAPTIC ACTIVE ZONES.
CC -! SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC -! SIMILARITY: BELONGS TO THE SYNAXIN/EPIMORPHIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D37933; BAA07152.1; -
CC HSP; P32851; 1BR0.
CC MIM; 601485; -
CC InterPro: IPR000017; SynTaxin.
CC DR InterPro: IPR000727; T-SNARE.
CC Pfam; PF00804; SynTaxin; 1.
CC SMART; SM00503; SYN; 1.
CC DR SMART; SM00397; T-SNARE; 1.
CC DR PROSITE; PS00914; SYNAXIN; 1.
CC KW Neurotransmitter transport; Coiled coil; Transmembrane.
CC FT DOMAIN 29 104 COILED COIL (POTENTIAL).
CC FT TRANSMEM 265 288 POTENTIAL.
CC SQ SEQUENCE 288 AA; 33431 MW; EC88271F01B1448A CRC64;

```

Query Match

Best Local Similarity 49.5%; Score 46; DB 1; Length 288;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

OY 1 LRRFRNKKIKKKIGOKI 19
   ::||::||| | |
DB 81 IKKTANKVRSKLEKAEOSI 99

```

Search completed: July 12, 2002, 08:04:43
Job time: 655 sec

Db 206 LKKLOERVKEOEKVEKEI 224

RESULT 6
Q9BPM9 PRELIMINARY; PRT: 523 AA.

AC 09BPM9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHEtical 58.7 KDA PROTEIN.
GN Y32G9A.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Bemis G., Courtney L.;
RT "The sequence of C. elegans cosmid Y32G9A.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR EMBL: AC087794; AAG53705.1; -
DR HSSP: P01008; IATH.
DR InterPro: IPR000215; Serpin.
DR Pfam: PF00079; serpin; 1.
DR SMART: SM00093; Serpin; 1.
DR Hypothetical protein; Serpin.
KW SEQUENCE 523 AA; 58718 MW; CAB3099CE357A199 CRC64;
SQ

Query Match 55.9%; Score 52; DB 5; Length 523;
Best Local Similarity 55.6%; Pred. No. 35;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 LKFRNKIKKIKKIGOK 18
Db 225 LPKLRNLSKLEALKLNEK 242

RESULT 7
Q06681 PRELIMINARY; PRT: 142 AA.

AC 06681;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHEtical 16.4 KDA PROTEIN.
GN AQ_635.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjey M., Huber R.,

RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
DR EMBL: AE000699; AAC06842.1; -
DR InterPro: IPR000106; Low_mwt_PTPase.
DR Pfam: PF01451; LMWpc; 1.
DR SMART: SM00226; LMWpc; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 142 AA; 16403 MW; C3AF01A2783FF145 CRC64;
SQ

Query Match 52.7%; Score 49; DB 16; Length 142;
Best Local Similarity 60.0%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 1 LKFRNKIKKIKKIKK 15
Db 115 LRRIRDKIEELKKL 129

RESULT 8
Q42555 PRELIMINARY; PRT: 331 AA.

AC Q42555;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CTP:PHOSPHOCHOLINE CYTIDYLTRANSFERASE (Ec 2.7.7.15).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RA Choi S.-B., Lee K.-W., Cho S.H.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U50451; AAA93035.1; -
DR HSSP: P27623; ICOZ.
DR InterPro: IPR001994; Cytidylyltransf.
DR Pfam: PF01467; Cytidylyltransf. 1.
KW Transferase; Nucleotidyltransferase.
SQ SEQUENCE 331 AA; 38417 MW; 85D7B9594E4E0B04 CRC64;
SQ

Query Match 52.7%; Score 49; DB 10; Length 331;
Best Local Similarity 36.8%; Pred. No. 57;
Matches 7; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

OY 1 LKFRNKIKKIKKIGOKI 19
Db 206 LKKLOERVKEOEKVEKEI 224

RESULT 9
Q9ZV56 PRELIMINARY; PRT: 332 AA.

AC Q9ZV56;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE PHOSPHOCHOLINE CYTIDYLTRANSFERASE.
GN T32F6.22 OR CCF.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;

RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.,
RT "Arabidopsis thaliana chromosome II BAC T3F6 genomic sequence."
RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP
RP STRAIN-CV, COLUMBIA COL-0;
RC
RA Choi Y.H., Choi S.B., Cho S.H.;
RT "Structure of a CTP:Phosphocholine Cytidyllyltransferase Gene from
RT Arabidopsis thaliana."
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC005700; AAC69950.1; -
DR EMBL: AF165912; AAD45922.1; -
DR HSSP: P27623; IC02.
DR InterPro: IPR001994; Cytidyllyltransf.
DR Pfam: PF01467; Cytidyllyltransf. 1
DR Nucleotide: Arabidopsis thaliana CTP:Phosphocholine Cytidyllyltransferase.
QO SEQUENCE 332 AAC; 38485 MW; 53562PFB87907AB54 CRC64;
QO

Query Match	52.7%	Score 49;	DB 10;	Length 332;
Best Local Similarity	36.8%;	Pred. No. 57;		
Matches	7; Conservative	9; Mismatches	3; Indels	0; Gaps

QY	1	LKKFRNKIKKKLKKIGOKI	19
Db	207	LKKLOERVRKCOOERVRGEKI	225

RESULT	10	
089295		
ID	Q89295	PRELIMINARY; PRT; 400 AA

DT	01-NOV-1996	(TEMBLrel. 01, Created)
DT	01-NOV-1996	(TEMBLrel. 01, Last sequence update)
DT	01-DEC-2001	(TEMBLrel. 19, Last annotation update)
DE	POLYPROTEIN (FRAGMENT).	
OS	Yellow fever virus.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
OC	Flavivirus.	
NCBI_taxid=11089;		
NCBI	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN-BRAZIL35.	
RC		
RX	MEDLINE-97078788; PubMed-8918913;	
RA	Wang E., Weaver S.C., Shope R.E., Tesh R.B., Watts D.M., Barrett A.D.,	
RT	"Genetic variation in yellow fever virus: duplication in the 3'	
RT	noncoding region of strains from Africa."	
RL	Virology 325:274-281(1996).	
DR	EMBL, U52389; AAB01953.1..	
DR	InterPro: IPR001122; Flavi_Capsid.	
DR	InterPro: IPR000336; Flavi_glycoprote.	
DR	InterPro: IPR000069; Flavi_M.	
DR	InterPro: IPR002535; Flavi_prop.	
DR	Pfam: PF01003; Flavi_Capsid; 1.	
DR	Pfam: PF00869; Flavi_glycoprote; 1.	
DR	Pfam: PF01004; Flavi_M; 1.	
DR	Pfam: PF01570; Flavi_prop; 1.	
DR	Prodom: PD001556; Flavi_glycoprote; 1.	
FT	NON_TER	
FT	400	
FT	400	
SEQUENCE	400 AA: 44278 MM: ADAC0A8E44EF877C CRC64:	

Query Match	52.7%	Score 49;	DB 12;	Length 400;
Best Local Similarity	50.0%;	Pred. No. 68;		
Matches	9;	Conservative	4;	Mismatches 5;
				Indels 0;
				Gaps 0

```
QY      1 LKRFNKIKEKLKIGOK 18
          :| |||:| |:||:
Db      20 VRSLSNKIKQKTKQIGNR 37
```

RESULT 11

ID	Q89297	PRELIMINARY:	PRT:	400 AA.
AC	Q89297;			
DT	01-NOV-1996	(TREMBLrel, 01, Created)		
DT	01-NOV-1996	(TREMBLrel, 01, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel, 19, Last annotation update)		
DE	POLYPROTEIN (FRAGMENT).			
OS	Yellow fever virus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;			
OC	Flavivirus.			
OX	NCBI_TaxID=11089;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CAR77(883);			
RX	MEDLINE=9/076788; PubMed=6918913;			
RA	Wang E., Weaver S.C., Shope R.E., Tesh R.B., Watts D.M., Barrett A.D.;			
RT	"Genetic variation in yellow fever virus: duplication in the 3'			
RT	noncoding region of strains from Africa.";			
RL	Virology 225:274-281(1996).			
DR	EMBL: U52392; AAB01955.1; -;			
DR	InterPro: IPR001122; Flavi_capsid			
DR	InterPro: IPR0000336; Flavi_glycoprote.			
DR	InterPro: IPR000069; Flavi_M.			
DR	InterPro: IPR002535; Flavi_prop.			
DR	Pfam: PF01003; Flavi_capsid; 1.			
DR	Pfam: PF00869; Flavi_glycoprot; 1.			
DR	Pfam: PF01004; Flavi_M; 1.			
DR	Pfam: PF01570; Flavi_prop; 1.			
DR	ProDom: PD001556; Flavi_glycoprote; 1.			
FT	NON_TER 400 400			
SQ	SEQUENCE 400 AA; 44491 MW; 6919JA40FEF74948 CRC64;			

Query Match	52.7%;	Score 49;	DB 12;	length 400;
Best Local Similarity	50.0%;	Pred. No. 68;		
Matches	9;	Conservative	4;	Mismatches 5;
				Indels 0;
				Gaps 0;

```
QY      1 LKKFRNKIKEKLLKIGOK 18
        :| |||:| |:|| :
Db      20 VRSLSNKIKQKTKQIGNR 37
```

RESULT	12		
Q89299			
ID	Q89299	PRELIMINARY;	PRT, 400 AA.
AC	Q89299;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	POLYPROTEIN (FRAGMENT).		
OS	Yellow fever virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		
OC	Flavivirus.		
OX	NCBI_TaxID=11089;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CAR77(900);		
RX	MEDLINE=97076788; PubMed=8918913;		
RA	Wang E., Weaver S.C., Shope R.E., Tesh R.B., Watts D.M., Barrett A.D.;		
RT	"Genetic variation in yellow fever virus: duplication in the 3'		
RT	noncoding region of strains from Africa.";		
RL	Virology 225:274-281(1996).		
DR	EMBL; U52395; AAB01957.1; -		
DR	InterPro; IPR001122; Flavi1_capsid.		
DR	InterPro; IPR000336; Flavi1_glycoprote.		
DR	InterPro; IPR000069; Flavi1_M.		
DR	InterPro; IPR002535; Flavi1_propep.		
DR	Pfam; PF01003; Flavi1_capsid; 1.		
DR	Pfam; PF00869; Flavi1_glycoproc; 1.		
DR	Pfam; PF01004; Flavi1_M; 1.		
DR	Pfam; PF01570; Flavi1_propep; 1.		
DR	ProDom; PD001556; Flavi1_glycoprote; 1.		
FT	NON_TER 400 400		

```
SO SEQUENCE 400 AA: 44491 MW: 69191A0FE674948 CRC64;

Query Match
Best Local Similarity 52.7%; Score 49; DB 12; Length 400;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 LKFRNKIKKKIKGK 18
: | ||||: | ||:
DB 20 VRSLSNKIKKTKQIGNR 37

RESULT 13
O89304 PRELIMINARY: PRT; 400 AA.
AC 089304;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Yellow fever virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID-11089;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-NIGERIA46;
RX MEDLINE-97078788; PubMed-8918913;
RA Wang E., Weaver S.C., Shope R.E., Tesh R.B., Watts D.M., Barrett A.D.;
RT "Genetic variation in yellow fever virus: duplication in the 3'
noncoding region of strains from Africa.";
RL Virology 225:274-281(1996).
DR EMBL: U52403; AAB01962.1;
DR InterPro: IPR001122; Flavi_capsid.
DR InterPro: IPR000336; Flavi_glycoprote.
DR InterPro: IPR000069; Flavi_M.
DR InterPro: IPR002535; Flavi_propep.
DR Pfam: PF01003; Flavi_capsid; 1.
DR Pfam: PF00869; Flavi_glycoprot; 1.
DR Pfam: PF01004; Flavi_M; 1.
DR Pfam: PF01570; Flavi_propep; 1.
DR ProDom: PD001556; Flavi_glycoprote; 1.
DR NON_TER 400
FT SEQUENCE 400 AA: 44365 MW: 75F6A5858658D6C1 CRC64;

Query Match
Best Local Similarity 52.7%; Score 49; DB 12; Length 400;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 LKFRNKIKKKIKGK 18
: | ||||: | ||:
DB 20 VRSLSNKIKKTKQIGNR 37

RESULT 14
O89311 PRELIMINARY: PRT; 400 AA.
AC 089311;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Yellow fever virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID-11089;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-SENEGAL65;
RX MEDLINE-97078788; PubMed-8918913;
RA Wang E., Weaver S.C., Shope R.E., Tesh R.B., Watts D.M., Barrett A.D.;
RT "Genetic variation in yellow fever virus: duplication in the 3'
```

```
RT noncoding region of strains from Africa.";
RL Virology 225:274-281(1996).
DR EMBL: U52413; AAB01969.1;
DR InterPro: IPR001122; Flavi_capsid.
DR InterPro: IPR000336; Flavi_glycoprote.
DR InterPro: IPR000069; Flavi_M.
DR InterPro: IPR002535; Flavi_propep.
DR Pfam: PF01003; Flavi_capsid; 1.
DR Pfam: PF00869; Flavi_glycoprot; 1.
DR Pfam: PF01004; Flavi_M; 1.
DR Pfam: PF01570; Flavi_propep; 1.
DR ProDom: PD001556; Flavi_glycoprote; 1.
DR NON_TER 400
FT SEQUENCE 400 AA: 44425 MW: DA6D0DF069F11591 CRC64;

Query Match
Best Local Similarity 52.7%; Score 49; DB 12; Length 400;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 LKFRNKIKKKIKGK 18
: | ||||: | ||:
DB 20 VRSLSNKIKKTKQIGNR 37

RESULT 15
O89315 PRELIMINARY: PRT; 400 AA.
AC 089315;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Yellow fever virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID-11089;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-TRINIDAD79;
RX MEDLINE-97078788; PubMed-8918913;
RA Wang E., Weaver S.C., Shope R.E., Tesh R.B., Watts D.M., Barrett A.D.;
RT "Genetic variation in yellow fever virus: duplication in the 3'
noncoding region of strains from Africa.";
RL Virology 225:274-281(1996).
DR EMBL: U52419; AAB01973.1;
DR InterPro: IPR001122; Flavi_capsid.
DR InterPro: IPR000336; Flavi_glycoprote.
DR InterPro: IPR000069; Flavi_M.
DR InterPro: IPR002535; Flavi_propep.
DR Pfam: PF01003; Flavi_capsid; 1.
DR Pfam: PF00869; Flavi_glycoprot; 1.
DR Pfam: PF01004; Flavi_M; 1.
DR Pfam: PF01570; Flavi_propep; 1.
DR ProDom: PD001556; Flavi_glycoprote; 1.
DR NON_TER 400
FT SEQUENCE 400 AA: 44452 MW: F04E08CFC24F73E0 CRC64;

Query Match
Best Local Similarity 52.7%; Score 49; DB 12; Length 400;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 LKFRNKIKKKIKGK 18
: | ||||: | ||:
DB 20 VRSLSNKIKKTKQIGNR 37
```

Search completed: July 12, 2002, 08:03:53
Job time: 639 sec

Shah
642744

=> fil reg
COST IN U.S. DOLLARS

SINCE FILE
ENTRY
2.94

TOTAL
SESSION
2.94

FULL ESTIMATED COST

FILE 'REGISTRY' ENTERED AT 09:00:30 ON 12 JUL 2002
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2002 American Chemical Society (ACS)

STRUCTURE FILE UPDATES: 10 JUL 2002 HIGHEST RN 438186-75-5
DICTIONARY FILE UPDATES: 10 JUL 2002 HIGHEST RN 438186-75-5

TSCA INFORMATION NOW CURRENT THROUGH January 7, 2002

Please note that search-term pricing does apply when
conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Calculated physical property data is now available. See HELP PROPERTIES
for more information. See STNote 27, Searching Properties in the CAS
Registry File, for complete details:
<http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf>

=> e knlrrriirkiihiikkyg/sqep

E1	1	KNLRRRIIRKIEHIIKKYG/SQEP
E2	1	KNLRRRIIRKIGHIIKKYG/SQEP
E3	2 -->	KNLRRRIIRKIIHIKKYG/SQEP
E4	1	KNLRRRIIRKIIHIKKYGPTILRIIRIIG/SQEP
E5	1	KNLRRRIIRKISHIIKKYG/SQEP
E6	1	KNLRRRIIRKITHIIKKYG/SQEP
E7	1	KNLRRRIIRKSIHIIKKYG/SQEP
E8	1	KNLRRRIIRKTIHIIKKYG/SQEP
E9	1	KNLRRISRKIIHIKKYG/SQEP
E10	1	KNLRRITRKIIHIKKYG/SQEP
E11	1	KNLRRSIRKIIHIKKYG/SQEP
E12	1	KNLRRTIRKIIHIKKYG/SQEP

=> s e3;d sqide cbib abs 1-2'e knirriirkiihiikkyg/sqep

2 KNLRRRIIRKIIHIKKYG/SQEP
63093 SQL=18

L1 2 (KNLRRRIIRKIIHIKKYG)/SQEP
(KNLRRRIIRKIIHIKKYG/SQEP AND SQL=18)

'KNIRRIIRKIIHIKKYG' MUST END IN '/Q', '/A', '/L', '/S' OR '/B'
The saved name for a query (or structure or screen set) must end with
'/Q'. The saved name for an answer set must end with '/A'. The saved
name for an L# list must end with '/L'. SDI request names must end
with '/S'. To see a list of all saved query, answer set,, and L# list
names for this loginid, enter "DISPLAY SAVED" at an arrow
prompt (=>). Enter "DISPLAY SAVED/S" to see a list of SDI request
names. Enter "DISPLAY SAVED/B" to see a list of BATCH search
requests.

=> s e3;d sqide cbib abs 1-2;e knirriirkiihiikkyg/sqep

2 KNLRRRIIRKIIHIKKYG/SQEP
63093 SQL=18

L2 2 (KNLRRRIIRKIIHIKKYG)/SQEP
(KNLRRRIIRKIIHIKKYG/SQEP AND SQL=18)

Searched by: Mary Hale 308-4258 CM-1 1E01

L2 ANSWER 1 OF 2 REGISTRY COPYRIGHT 2002 ACS
RN 326855-45-2 REGISTRY
CN Glycine, L-lysyl-L-asparaginyl-L-leucyl-L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-isoleucyl-L-isoleucyl-L-lysyl-L-lysyl-L-tyrosyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 8: PN: W00112668 SEQID: 8 claimed protein
CN Ovispirin OV 1
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 18

SEQ 1 KNLRRIIRKI IHIIKKYG
=====

HITS AT: 1-18

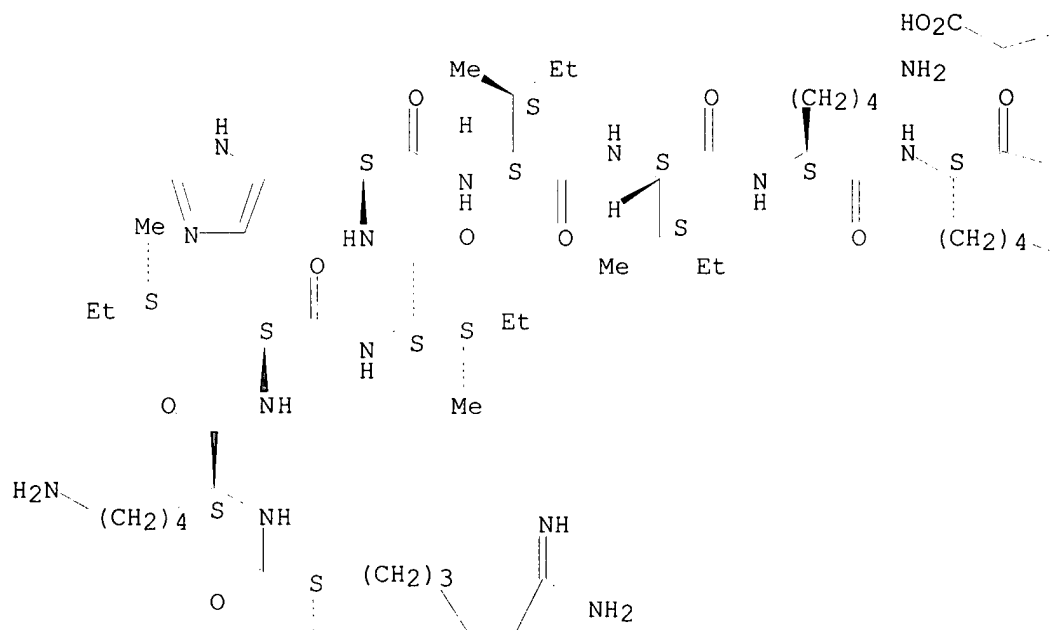
MF C105 H188 N34 O21

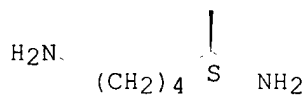
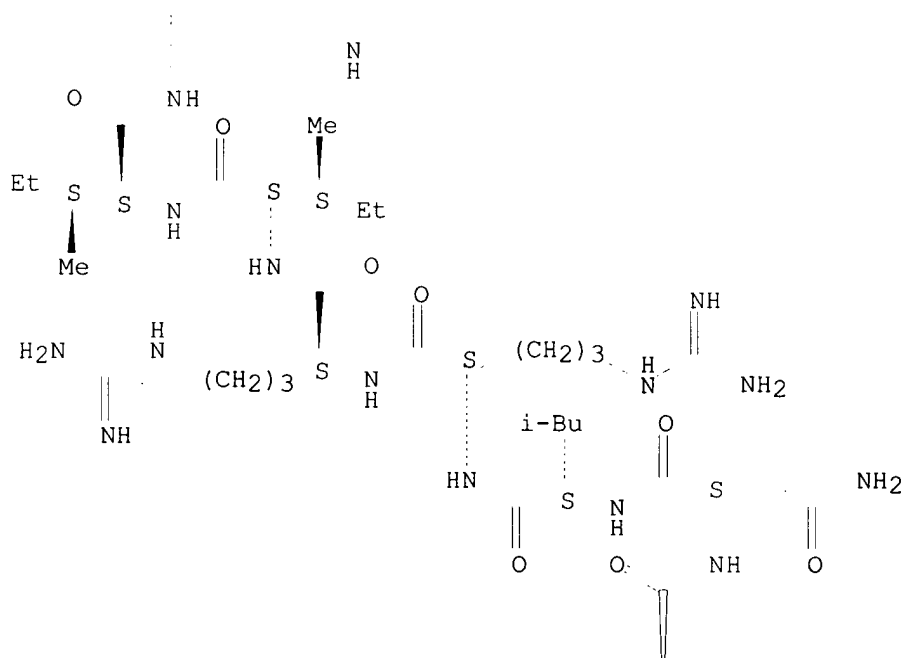
SR CA

LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.

PAGE 1-A





5 REFERENCES IN FILE CA (1967 TO DATE)
5 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for

Microbiology.

- AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PAO1.

REFERENCE 2: 136:74662 Pharmaceutical composition comprising novispirin. Lehrer, Robert I.; Waring, Alan J.; Tack, Brian F. (The Regents of the University of California, USA; The University of Iowa). PCT Int. Appl. WO 2002000839 A2 20020103, 42 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG, TR. (English). CODEN: PIXXD2. APPLICATION: WO 2001-US19094 20010613. PRIORITY: US 2000-606858 20000628; US 2001-840009 20010419.

- AB Novispririn peptides are antimicrobial agents with potent activity against Gram-neg. bacteria. The peptides are nonhemolytic, exhibit reduced in vitro cytotoxicity relative to other antimicrobial peptides, and were well-tolerated in vivo after i.v. injection. Novispirins also bind lipopolysaccharide (LPS), a property that may mitigate symptoms assocd. with Gram-neg. bacterial infection. A pharmaceutical compn. comprising novispirin as an active agent is administered to a patient suffering from or predisposed to a microbial infection, particularly Gram-neg. bacterial infections.

REFERENCE 3: 136:50035 Orientation and dynamics of an antimicrobial peptide in the lipid bilayer by solid-state NMR spectroscopy. Yamaguchi, Satoru; Huster, Daniel; Waring, Alan; Lehrer, Robert I.; Kearney, William; Tack, Brian F.; Hong, Mei (Department of Chemistry, Iowa State University, Ames, IA, 500112, USA). Biophysical Journal, 81(4), 2203-2214 (English) 2001. CODEN: BIOJAU. ISSN: 0006-3495. Publisher: Biophysical Society.

- AB The orientation and dynamics of an 18-residue antimicrobial peptide, ovispirin, has been investigated using solid-state NMR spectroscopy. Ovispirin is a cathelicidin-like model peptide (NH₂-KNLRRRIIRKIIHIKKYG-COOH) with potent, broad-spectrum bactericidal activity. 15N NMR spectra of oriented ovispirin reconstituted into synthetic phospholipids show that the helical peptide is predominantly oriented in the plane of the lipid bilayer, except for a small portion of the helix, possibly at the C-terminus, which deviates from the surface orientation. This suggests differential insertion of the peptide backbone into the lipid bilayer. 15N spectra of both oriented and unoriented peptides show a reduced 15N chem. shift anisotropy at room temp. compared with that of rigid proteins, indicating that the peptide undergoes uniaxial rotational diffusion around the bilayer normal with correlation times shorter than 10⁻⁴ s. This motion is frozen below the gel-to-liq. cryst. transition temp. of the lipids. Ovispirin interacts strongly with the lipid bilayer, as manifested by the significantly reduced 2H quadrupolar splittings of perdeuterated palmitoyl-oleoylphosphatidylcholine acyl chains upon peptide binding. Therefore, ovispirin is a curved helix residing in the membrane-water interface that executes rapid uniaxial rotation. These structural and dynamic features are important for understanding the antimicrobial function of this peptide.

REFERENCE 4: 136:2776 Cathelicidin peptides inhibit multiply antibiotic-resistant pathogens from patients with cystic fibrosis. Saiman, Lisa; Tabibi, Setareh; Starner, Timothy D.; San Gabriel, Pablo;

Winokur, Patricia L.; Jia, Hong Peng; McCray, Paul B., Jr.; Tack, Brian F. (Department of Pediatrics, Columbia University, New York, NY, 10032, USA). Antimicrobial Agents and Chemotherapy, 45(10), 2838-2844 (English) 2001. CODEN: AMACQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB Endogenous peptide antibiotics are under investigation as inhaled therapeutic agents for cystic fibrosis (CF) lung disease. The bactericidal activities of five cathelicidin peptides (LL37 [human], CAP18 [rabbit], mCRAMP [mouse], rCRAMP [rat], and SMAP29 [sheep]), three novel alpha-helical peptides derived from SMAP29 and termed ovispirins (OV-1, OV-2, and OV-3), and two derivs. of CAP18 were tested by broth microdilution assays. Their MICs were detd. for multiply antibiotic-resistant *Pseudomonas aeruginosa* (n = 24), *Burkholderia cepacia* (n = 5), *Achromobacter xylosoxidans* (n = 5), and *Stenotrophomonas maltophilia* (n = 5) strains isolated from CF patients. SMAP29 was most active and inhibited mucoid and nonmucoid *P. aeruginosa* strains (MIC, 0.06 to 8 .mu.g/mL). OV-1, OV-2, and OV-3 were nearly as active (MIC, 0.03 to 16 .mu.g/mL), but CAP18 (MIC, 1.0 to 32 .mu.g/mL), CAP18-18 (MIC, 1.0 to >32 .mu.g/mL), and CAP18-22 (MIC, 0.5 to 32 .mu.g/mL) had variable activities. LL37, mCRAMP, and rCRAMP were least active against the clin. isolates studied (MIC, 1.0 to >32 .mu.g/mL). Peptides had modest activities against *S. maltophilia* and *A. xylosoxidans* (MIC range, 1.0 to > 32 .mu.g/mL), but none inhibited *B. cepacia*. However, CF sputum inhibited the activity of SMAP29 substantially. The effects of peptides on bacterial cell membranes and eukaryotic cells were examd. by SEM and by measuring transepithelial cell resistance, resp. SMAP29 caused the appearance of bacterial membrane blebs within 1 min, killed *P. aeruginosa* within 1 h, and caused a dose-dependent, reversible decrease in transepithelial resistance within 5 h. The tested cathelicidin-derived peptides represent a novel class of antimicrobial agents and warrant further development as prophylactic or therapeutic agents for CF lung disease.

REFERENCE 5: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

L2 ANSWER 2 OF 2 REGISTRY COPYRIGHT 2002 ACS

RN 326855-39-4 REGISTRY

CN Glycinamide, L-lysyl-L-asparaginyl-L-leucyl-L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-

L-isoleucyl-L-isoleucyl-L-lysyl-L-lysyl-L-tyrosyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 1: PN: WO0112668 SEQID: 1 claimed protein

CN Ovispirin OV 3

FS PROTEIN SEQUENCE; STEREOSEARCH

SQL 18

NTE modified

type	----- location -----	description
terminal mod.	Gly-18 -	C-terminal amide

SEQ 1 KNLRRRIIRKI IHIIKKYG

=====

HITS AT: 1-18

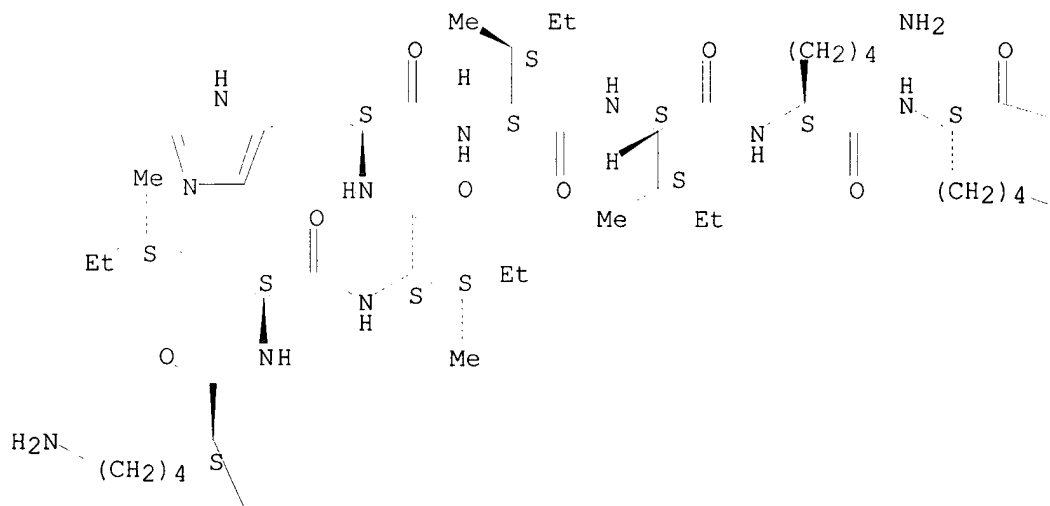
MF C105 H189 N35 O20

SR CA

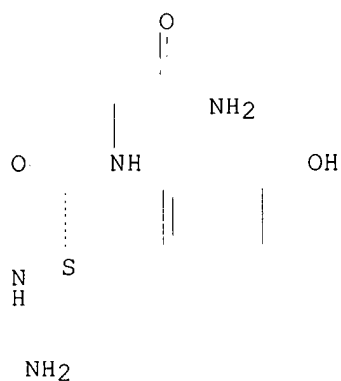
LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.

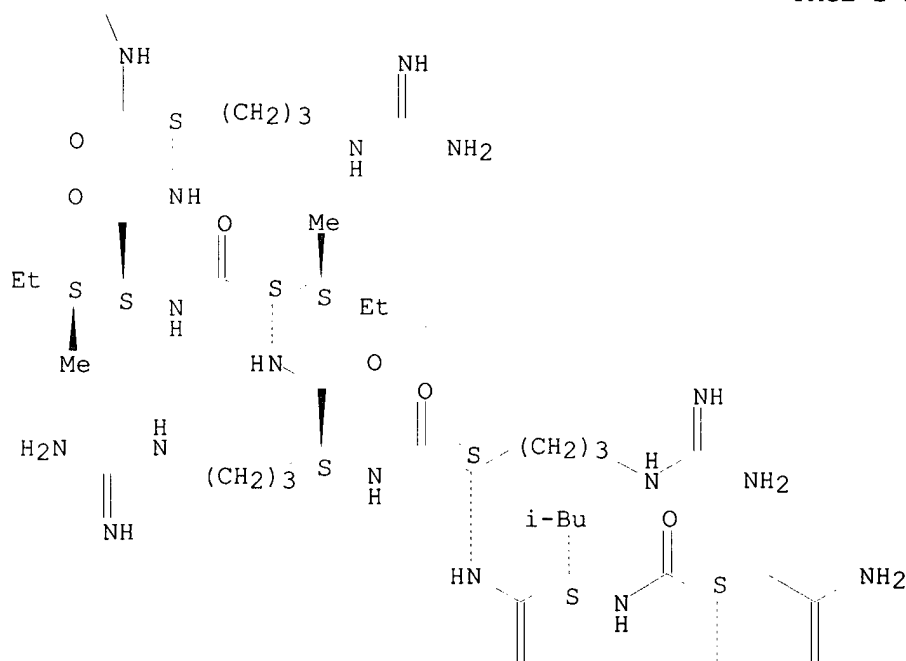
PAGE 1-A



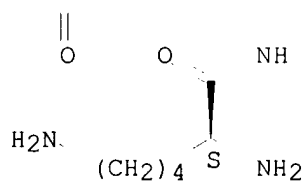
PAGE 1-B



PAGE 2-A



PAGE 3-A



3 REFERENCES IN FILE CA (1967 TO DATE)
3 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce

Searched by: Mary Hale 308-4258 CM-1 1E01

ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

- AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PA01.

REFERENCE 2: 136:74662 Pharmaceutical composition comprising novispirin. Lehrer, Robert I.; Waring, Alan J.; Tack, Brian F. (The Regents of the University of California, USA; The University of Iowa). PCT Int. Appl. WO 2002000839 A2 20020103, 42 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG, TR. (English). CODEN: PIXXD2. APPLICATION: WO 2001-US19094 20010613. PRIORITY: US 2000-606858 20000628; US 2001-840009 20010419.

- AB Novispirin peptides are antimicrobial agents with potent activity against Gram-neg. bacteria. The peptides are nonhemolytic, exhibit reduced in vitro cytotoxicity relative to other antimicrobial peptides, and were well-tolerated in vivo after i.v. injection. Novispirins also bind lipopolysaccharide (LPS), a property that may mitigate symptoms assocd. with Gram-neg. bacterial infection. A pharmaceutical compn. comprising novispirin as an active agent is administered to a patient suffering from or predisposed to a microbial infection, particularly Gram-neg. bacterial infections.

REFERENCE 3: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

- AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

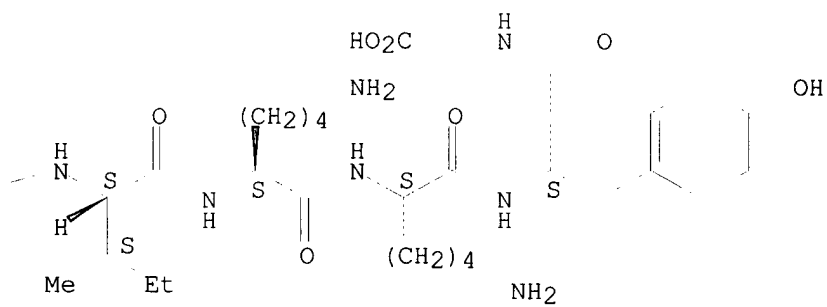
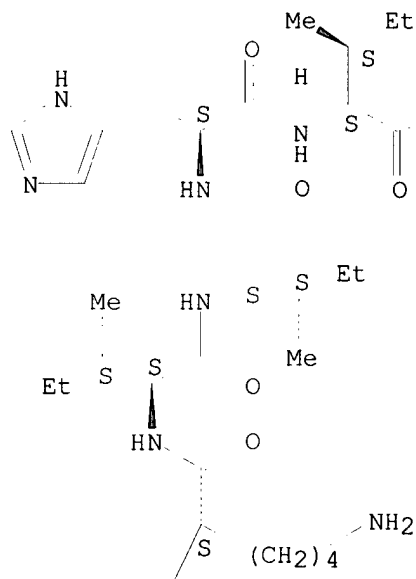
E1 1 KNIRELWPGAVAHACNPSTPGGRGRWITRSGDWLNLRTCEHVILQQKGKFRFLIS/SQEP
 E2 1 KNIRGILESLMCNESSIRNLRERKSVN/SQEP
 E3 2 --> KNIRRIIRKIIHIKKYG/SQEP
 E4 1 KNIRRIIRKIIHIKKYG/SQEP
 E5 1 KNIRRRVYDALNVLMAMNIISKEKKEIKWIGLPTNSA/SQEP
 E6 1 KNIRRRVYDGGNVLMAMNIISKEKKEIKWIGLPTNSA/SQEP
 E7 1 KNIRRSTLGEDKVEV/SQEP
 E8 1 KNISFFPSNVEQFSEGNIDVWWIVHDGGM/L/SQEP
 E9 3 KNISILDSQVIVTGNLDISTELGNVNNSISNALDK/SQEP
 E10 1 KNISQVSQSCGYNSTSYFISVFKDFYGMTPL/SQEP
 E11 1 KNISRAKWNNTLKQIDSKLREQF/SQEP
 E12 1 KNISRGRGGRYHASCVRWLMRALSLLASNQTKIGAAEQNDP/SQEP

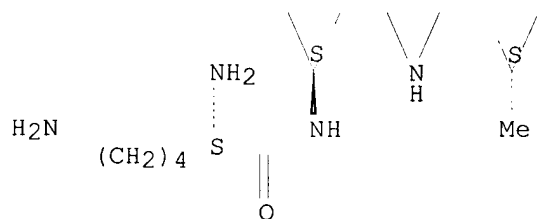
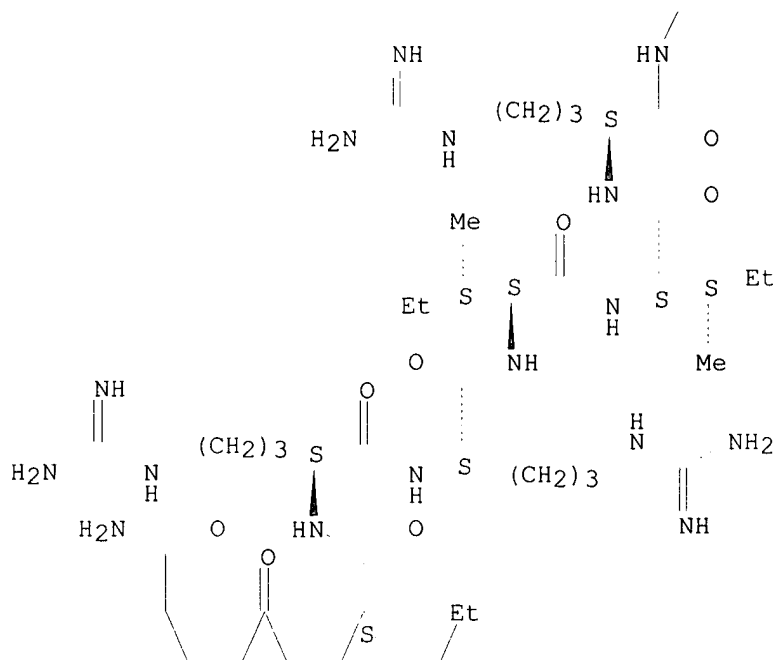
=> s e3;d sqide cbib abs 1-2;e nlrrriirkiihiikky/sqep
 2 KNIRRIIRKIIHIKKYG/SQEP
 63093 SQL=18
 L3 2 (KNIRRIIRKIIHIKKYG)/SQEP
 (KNIRRIIRKIIHIKKYG/SQEP AND SQL=18)

L3 ANSWER 1 OF 2 REGISTRY COPYRIGHT 2002 ACS
 RN 386702-96-1 REGISTRY
 CN Glycine, L-lysyl-L-asparaginyl-L-isoleucyl-L-arginyl-L-arginyl-L-isoleucyl-
 L-isoleucyl-L-arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-
 isoleucyl-L-isoleucyl-L-lysyl-L-lysyl-L-tyrosyl- (9CI) (CA INDEX NAME)
 OTHER NAMES:
 CN Ovispirin OV 5
 FS PROTEIN SEQUENCE; STEREOSEARCH
 SQL 18

SEQ 1 KNIRRIIRKI IHIKKYG
 =====
 HITS AT: 1-18
 MF C105 H188 N34 O21
 SR CA
 LC STN Files: CA, CAPLUS

Absolute stereochemistry.





PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT

1 REFERENCES IN FILE CA (1967 TO DATE)

1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PAO1.

L3 ANSWER 2 OF 2 REGISTRY COPYRIGHT 2002 ACS

RN 326855-43-0 REGISTRY

CN Glycinamide, L-lysyl-L-asparaginyl-L-isoleucyl-L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-isoleucyl-L-isoleucyl-L-lysyl-L-lysyl-L-tyrosyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 6: PN: WO0112668 SEQID: 6 claimed protein

CN Ovispirin OV 4

FS PROTEIN SEQUENCE; STEREOSEARCH

SQL 18

NTE modified

type	location	description
terminal mod.	Gly-18	- C-terminal amide

SEQ 1 KNIRRIIRKI IHIIKKYG

HITS AT: 1-18

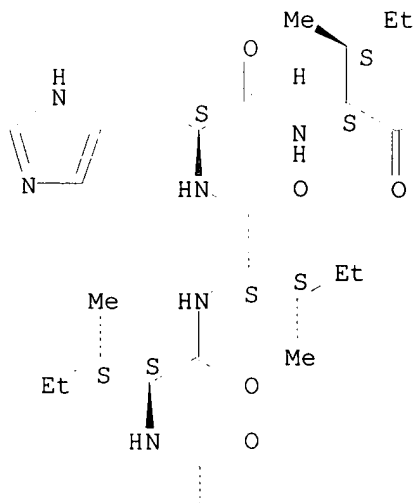
MF C105 H189 N35 O20

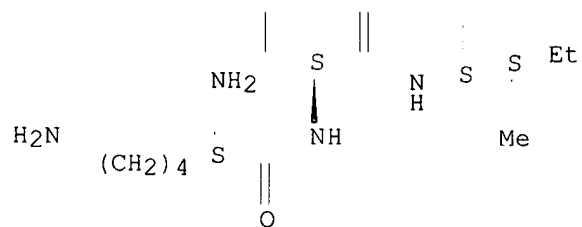
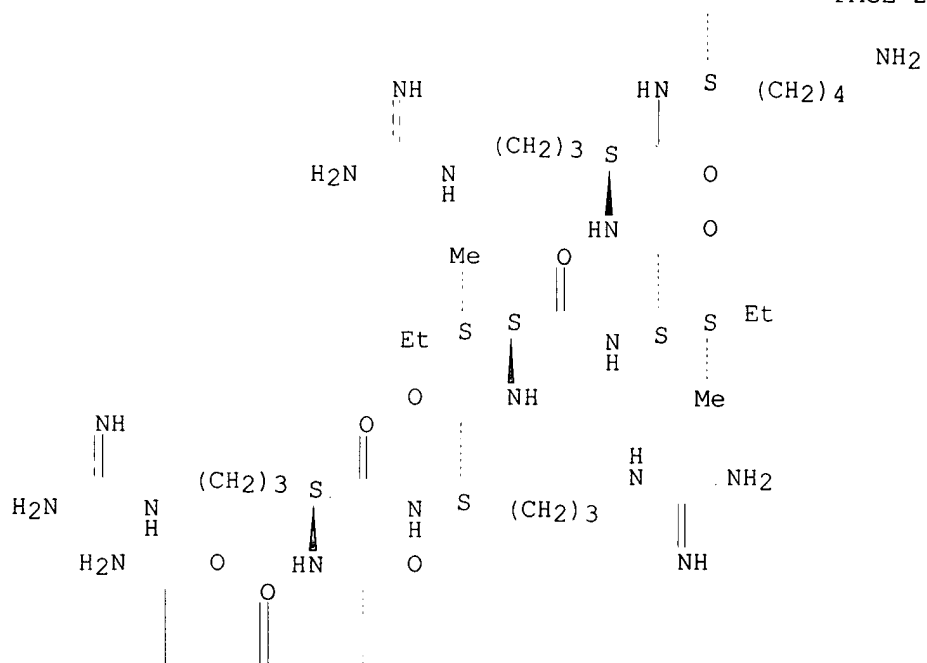
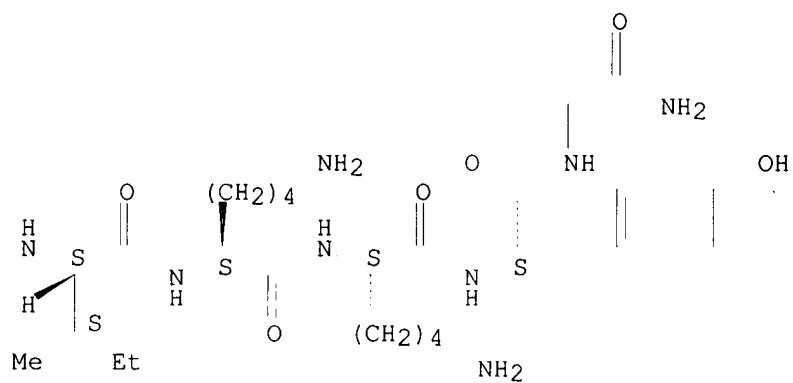
SR CA

LC STN Files: CA, CAPLUS

Absolute stereochemistry.

PAGE 1-A





2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of Pseudomonas aeruginosa PAO1.

REFERENCE 2: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

E1	1	NLRRGTALA/SQEP
E2	1	NLRRGTALAIPOCRL/SQEP
E3	2 -->	NLRRIRKIIHIKKY/SQEP
E4	1	NLRRLLLSHI/SQEP
E5	1	NLRRLLLSMI/SQEP
E6	1	NLRSAPWVGQDIVEFIWGGFSVNNATLNRFFALHFVLPFVLAALALMHLLIALHDSSGSG NPLGVSGNYDRLPFWPYFIFKDLITIFLFFIVLSIFVFFMPNVLDSENYVMANPMQTPP AIVPEWYLLPFYA/SQEP
E7	2	NLRSALKTHPAESTWYSWVQF/SQEP
E8	1	NLRSFIHKV/SQEP
E9	1	NLRSFY/SQEP
E10	1	NLRSPNNFL/SQEP
E11	1	NLRSRHCIRIHRFLYEEQPVPSWPSPPTRIDQDLPSACSPGGQRTPLPGEIRRAVQQGGT STISSVRHSSWAPWRRPITRRAPSLSTGTPTRTSAGRCSLLRSS/SQEP
E12	1	NLRSSRRVEDTWTLTQTPRWAVPRDPAREGIPSLRFSVLLWLEDLNPTSPSSGDRAF' AAA'RAGDLAVLRTVEQGGRGGLGWW'AAA'RGNRVRCVKQS/SQEP

=> s e3;d sqide cbib abs 1-2;e lrrriirkiihiikk/sqep
2 NLRRIRKIIHIKKY/SQEP

29730 SQL=16
 L4 2 (NLRRRIIRKIIHIIKKY)/SQEP
 (NLRRRIIRKIIHIIKKY/SQEP AND SQL=16)

L4 ANSWER 1 OF 2 REGISTRY COPYRIGHT 2002 ACS
 RN 326855-57-6 REGISTRY
 CN L-Tyrosinamide, L-asparaginyl-L-leucyl-L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-isoleucyl-L-isoleucyl-L-lysyl-L-lysyl- (9CI) (CA INDEX NAME)
 FS PROTEIN SEQUENCE; STEREOSEARCH
 SQL 16
 NTE modified

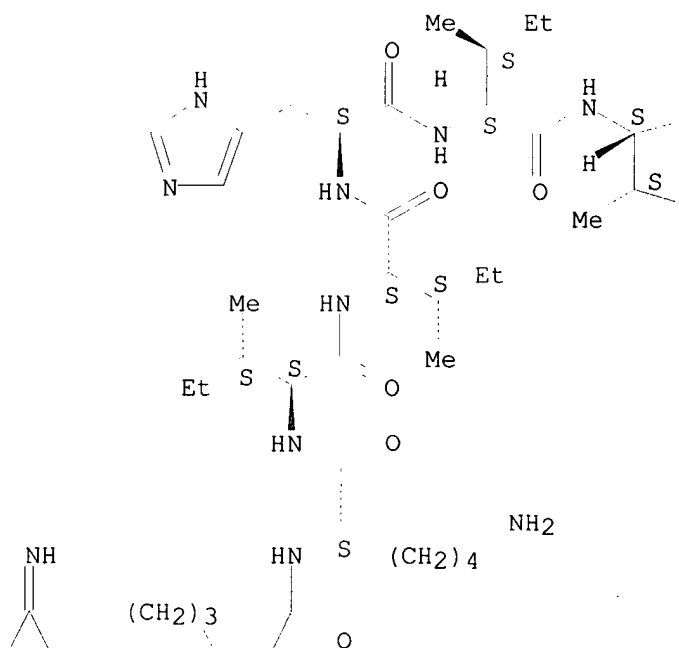
type	location	description
terminal mod.	Tyr-16	C-terminal amide

SEQ 1 NLRRRIIRKII HIIKKY
 =====

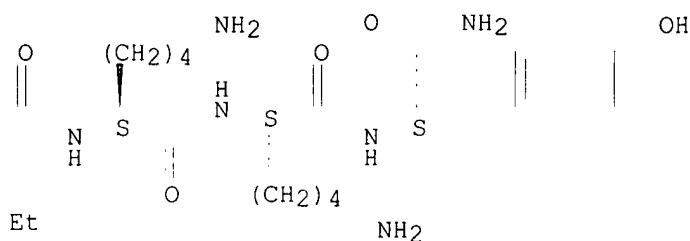
HITS AT: 1-16
 MF C97 H174 N32 O18
 SR CA
 LC STN Files: CA, CAPLUS

Absolute stereochemistry.

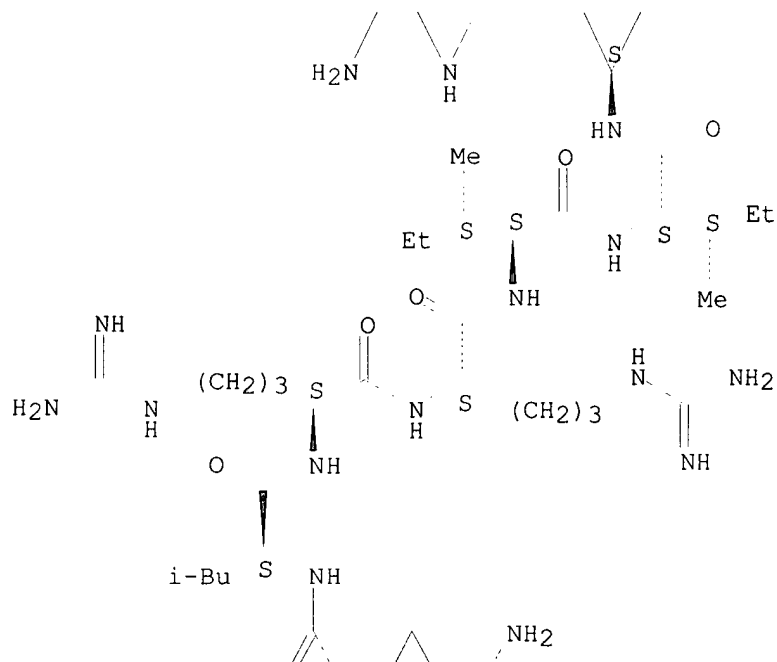
PAGE 1-A



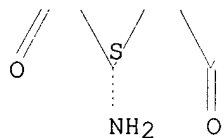
PAGE 1-B



PAGE 2-A



PAGE 3-A



1 REFERENCES IN FILE CA (1967 TO DATE)
1 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP,

Searched by: Mary Hale 308-4258 CM-1 1E01

KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2.

APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

L4 ANSWER 2 OF 2 REGISTRY COPYRIGHT 2002 ACS

RN 326855-46-3 REGISTRY

CN L-Tyrosine, L-asparaginyl-L-leucyl-L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-isoleucyl-L-isoleucyl-L-lysyl-L-lysyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 9: PN: WO0112668 SEQID: 9 claimed protein

CN Ovispirin OV 6

FS PROTEIN SEQUENCE; STEREOSEARCH

SQL 16

SEQ 1 NLRRIIRKII HIIKKY

=====

HITS AT: 1-16

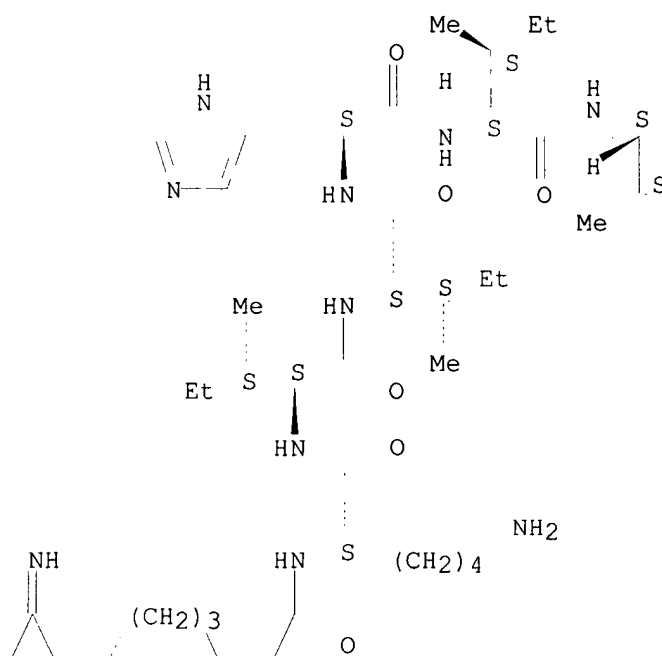
MF C97 H173 N31 O19

SR CA

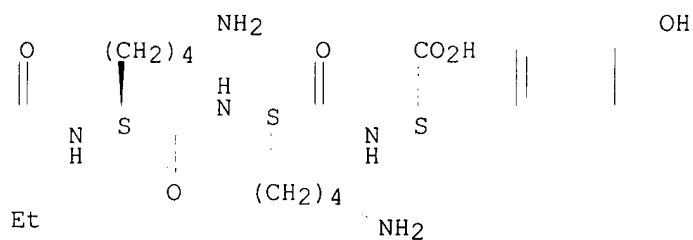
LC STN Files: CA, CAPLUS

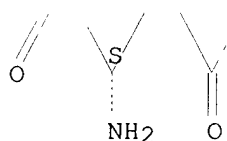
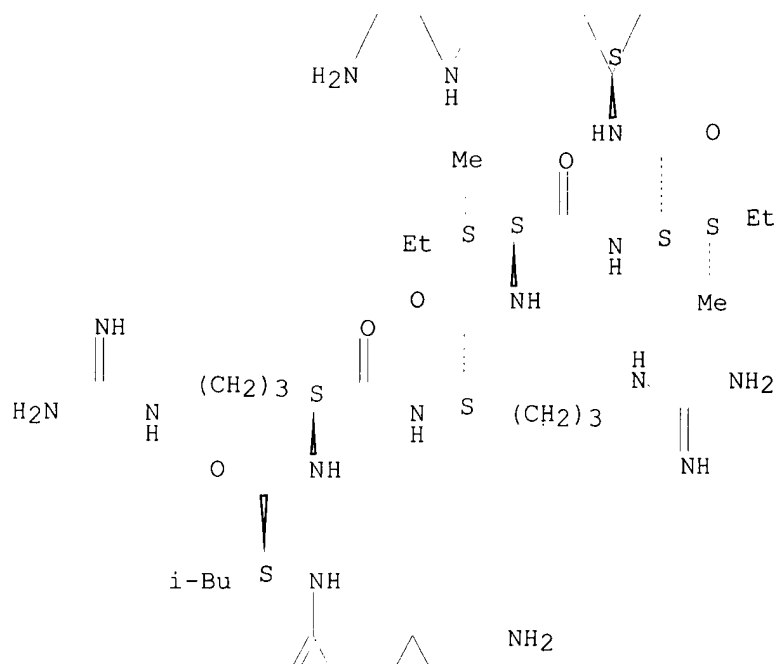
Absolute stereochemistry.

PAGE 1-A



PAGE 1-B





2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

- REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.
- AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PAO1.
- REFERENCE 2: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ,

EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2.
APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

```
E1      1      LRRIGDELD/SQEP
E2      1      LRRIRKIIHIK/SQEP
E3      2 --> LRRIRKIIHIKK/SQEP
E4      1      LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSAAYFSNFLQQ
          TKGKPPSWN'AAA'PSPFSLAFDPTKGL/SQEP
E5      1      LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSAAYFSNFLQQ
          TKGKPPSWNVSPFSLAFDPTKGL/SQEP
E6      1      LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSGAYFFDFLQQ
          TKGKPPSWNVSPFSLAFDP'AAA-AAA'GL/SQEP
E7      1      LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSGAYFFDFLQQ
          TKGKPPSWNVSPFSLAFDPTKGL/SQEP
E8      1      LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSGAYFSDFLQQ
          TKGKPPSWNVSPFSLAFDPT'AAA'GL/SQEP
E9      1      LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSGAYFSDFLQQ
          TKGKPPSWNVSPFSLAFDPTKGL/SQEP
E10     1      LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSGAYFSDFLQQ
          TKGNPSSLNVSPFSLAFDPTKG'AAA'/SQEP
E11     3      LRRIR/SQEP
E12     1      LRRIREVLRTELTY/SQEP
```

```
=> s e2;d sqide cbib abs 1-2;e irriirkiihiikk/sqep
          1 LRRIRKIIHIK/SQEP
          292001 SQL=13
L5      1 (LRRIRKIIHIK)/SQEP
          (LRRIRKIIHIK/SQEP AND SQL=13)
```

```
L5      ANSWER 1 OF 1  REGISTRY  COPYRIGHT 2002 ACS
RN      326855-52-1  REGISTRY
CN      L-Lysinamide, L-leucyl-L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-
          arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-isoleucyl-L-isoleucyl-
          (9CI) (CA INDEX NAME)
OTHER NAMES:
CN      15: PN: WO0112668 SEQID: 15 claimed protein
CN      Ovispirin OV 11
FS      PROTEIN SEQUENCE; STEREOSEARCH
SQL     13
NTE     modified
```

```
-----
type          ----- location ----- description
```

Searched by: Mary Hale 308-4258 CM-1 1E01

terminal mod. Lys-13 - C-terminal amide

SEQ 1 LRRIRKIIH IIK

=====

HITS AT: 1-13

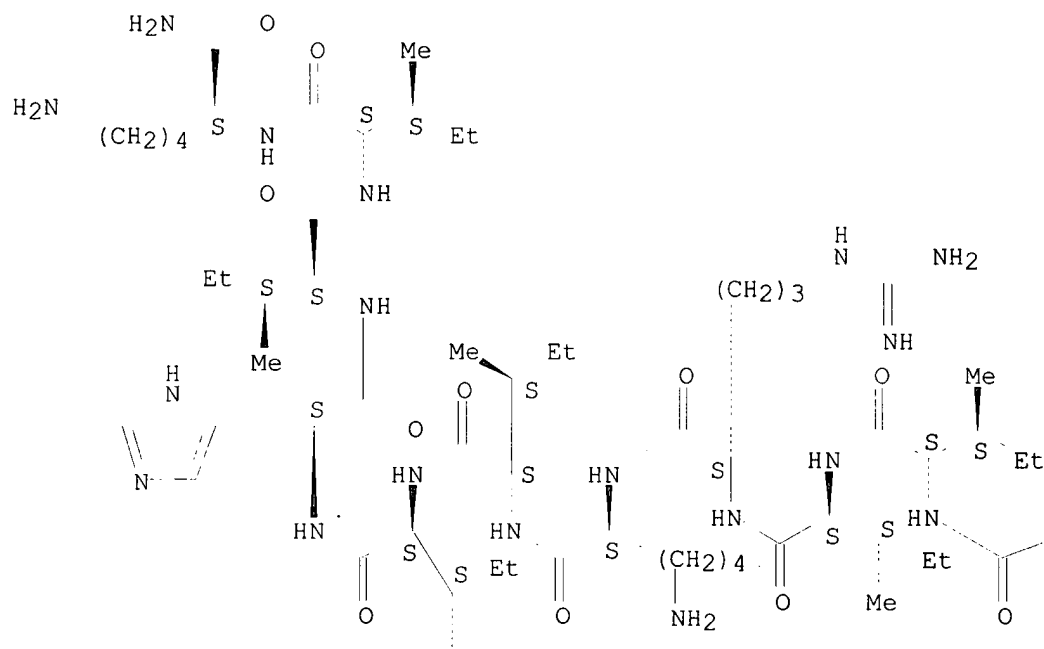
MF C78 H147 N27 O13

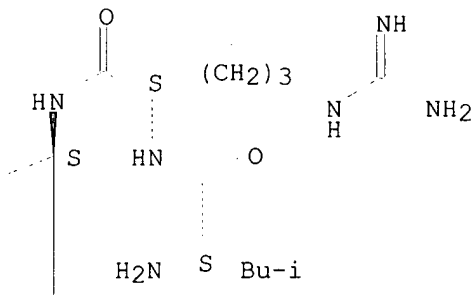
SR CA

LC STN Files: CA, CAPLUS

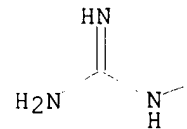
Absolute stereochemistry.

PAGE 1-A





Me



(CH₂)₃

2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

- REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.
- AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all

antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PAO1.

REFERENCE 2: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

```
E1      1      IRRHTGDKDDHPAKSWEPWQKRQIGYPGIGMAKHGVGPVITPSRNDGIPPSVPSSQPRND
          RSVRSHRRYEGQALSHPHGKTDADGDEPPTSVHILMTGRNRGCTHRDSMIALKVVLGRMA
          SATSSSSGW/SQEP
E2      1      IRRHKVIRP/SQEP
E3      2 --> IRRIRKIIHIKK/SQEP
E4      1      IRRINFTGSTRVGSIIAQKAAQHLKRCLLELGKKSPLIVLDDADIDAAVKAAVFGSFLFQ
          GQICMSTERLIVDEKIADEFVAKFVEKTKRLSAGDPCRLTGDCIIGPMVSPNSGERINGL
          FKDAIDKGAKVVCGLAQGALMPATMLDHVKSDMRIYDEETFGPITVVIRCKGEAEAVRI
          ANDSVYGLSSGVFGRDINRA/SQEP
E5      1      IRRIRMTNLTILTGRITKDLELKQAGQTQVTNFSMAVDNPFKKDDTSFFDIVAFGKTAQL
          LNDYCGKGSKVLIEGNLQDRFQDKEGHNRSVVRVIANRIEFLDSKGSNQNNQPQKQQE
          QTTEKSKNPFANASGPIDITDDMLPF/SQEP
E6      1      IRRK/SQEP
E7      1      IRRKEVN/SQEP
E8      2      IRRKFSPVL'AAA'MLMTSMNPA'AAA'VG'AAA'ISSDLAI'AAA'FNEPLHANLLYFI
          'AAA-AAA'QGILKSVPQENDEGETLS'AAA'L/SQEP
E9      1      IRRKIDMSKTAIITGAAGGLGKGIAERLANDGFNIVLQDINEALLLETEKEFKEKGYQAV
          AFKSDVSKKKEQEELVQFAVTEFGQLDVMVNNAGVDAVTPILEIGEEELSKLFNINVFGT
          LFGIQAAANQFIKQKSKGKIINACSIAGHESYEVLGTYSATKHSVRSFTQTAAKELADKG
          ITVNAYCPGVAKTEMWDRID/SQEP
E10     1      IRRKNGILHEDLDKEIYMTQHQGFSAPSQEHLVCHLQKSFYDLKQAPSQW'AAA'KRFD
          FMIAHGYYSNYDHCIYLKQCPNGSFVYLLLYVDDMDP/SQEP
E11     1      IRRKTEGT/SQEP
E12     1      IRRKTKEEKR'AAA'KEWREGREEFVENTFCLGTKRSQLSRNLKKLSIFGEFLECIFRFS
          RLLLI/SQEP
```

```
=> s e3;d sqide cbib abs 1-2;e lrrirkiihiik/sqep
      2 IRRIRKIIHIKK/SQEP
      32106 SQL=14
L6      2 (IRRIIRKIIHIKK)/SQEP
          (IRRIIRKIIHIKK/SQEP AND SQL=14)
```

L6 ANSWER 1 OF 2 REGISTRY COPYRIGHT 2002 ACS
RN 326855-51-0 REGISTRY
CN L-Lysine, L-isoleucyl-L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-
arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-isoleucyl-L-isoleucyl-
L-lysyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 14: PN: WO0112668 SEQID: 14 claimed protein
CN Ovispirin OV 10
FS PROTEIN SEQUENCE; STEREOSEARCH
SQL 14

SEQ 1 IRRIRKIIH IIKK
=====

HITS AT: 1-14

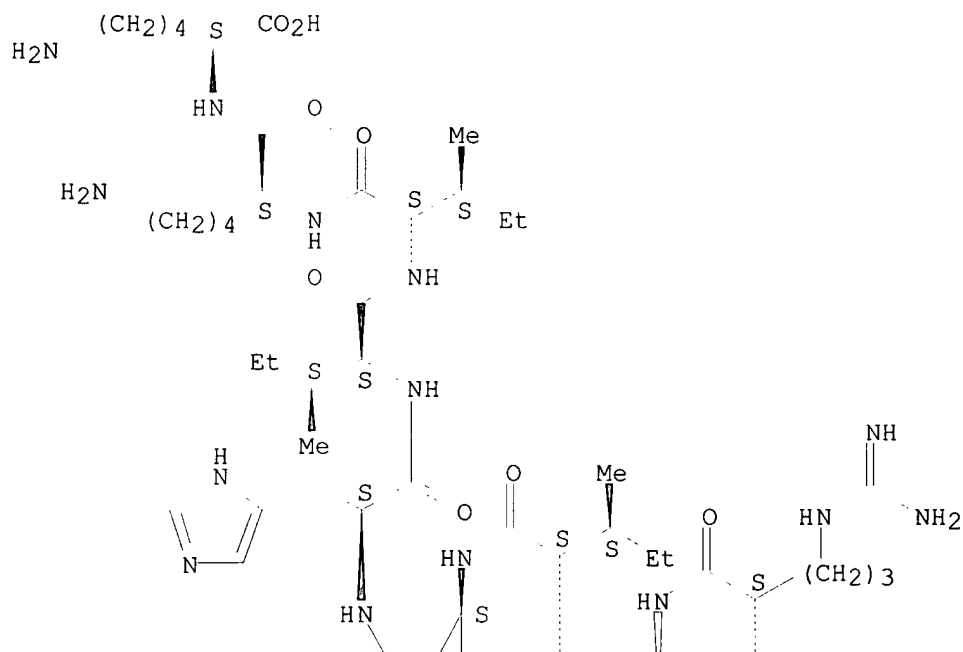
MF C84 H158 N28 O15

SR CA

LC STN Files: CA, CAPLUS

Absolute stereochemistry.

PAGE 1-A



REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PAO1.

REFERENCE 2: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ,

EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2.
APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

L6 ANSWER 2 OF 2 REGISTRY COPYRIGHT 2002 ACS

RN 326855-50-9 REGISTRY

CN L-Lysinamide, L-isoleucyl-L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-isoleucyl-L-isoleucyl-L-lysyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 13: PN: WO0112668 SEQID: 13 claimed protein

CN Ovispirin OV 9

FS PROTEIN SEQUENCE; STEREOSEARCH

SQL 14

NTE modified

type	----- location -----	description
terminal mod.	Lys-14 -	C-terminal amide

SEQ 1 IRRRIIRKIIH IIKK
=====

HITS AT: 1-14

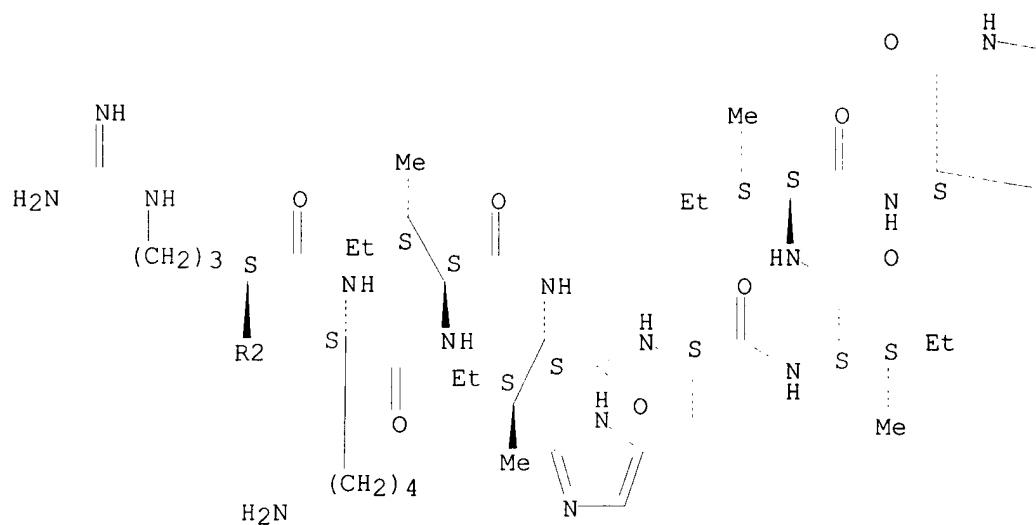
MF C84 H159 N29 O14

SR CA

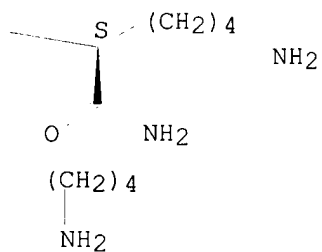
LC STN Files: CA, CAPLUS

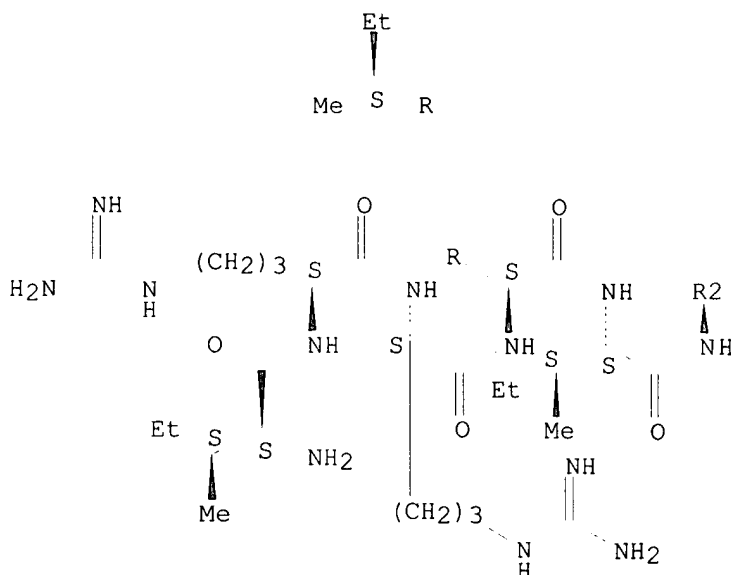
Absolute stereochemistry.

PAGE 1-A



PAGE 1-B





3 REFERENCES IN FILE CA (1967 TO DATE)
 3 REFERENCES IN FILE CAPLUS (1967 TO DATE)

- REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.
- AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PAO1.
- REFERENCE 2: 136:2776 Cathelicidin peptides inhibit multiply antibiotic-resistant pathogens from patients with cystic fibrosis. Saiman, Lisa; Tabibi, Setareh; Starner, Timothy D.; San Gabriel, Pablo; Winokur, Patricia L.; Jia, Hong Peng; McCray, Paul B., Jr.; Tack, Brian F. (Department of Pediatrics, Columbia University, New York, NY, 10032, USA). Antimicrobial Agents and Chemotherapy, 45(10), 2838-2844 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.
- AB Endogenous peptide antibiotics are under investigation as inhaled therapeutic agents for cystic fibrosis (CF) lung disease. The bactericidal activities of five cathelicidin peptides (LL37 [human], CAP18 [rabbit], mCRAMP [mouse], rCRAMP [rat], and SMAP29 [sheep]), three novel alpha-helical peptides derived from SMAP29 and termed ovispirins (OV-1, OV-2, and OV-3), and two derivs. of CAP18 were tested by broth microdilution assays. Their MICs were detd. for multiply antibiotic-resistant *Pseudomonas aeruginosa* (n = 24), *Burkholderia cepacia* (n = 5), *Achromobacter xylosoxidans* (n = 5), and *Stenotrophomonas maltophilia* (n = 5) strains isolated from CF patients. SMAP29 was most active and inhibited mucoid and nonmucoid *P. aeruginosa* strains (MIC, 0.06

to 8 .mu.g/mL). OV-1, OV-2, and OV-3 were nearly as active (MIC, 0.03 to 16 .mu.g/mL), but CAP18 (MIC, 1.0 to 32 .mu.g/mL), CAP18-18 (MIC, 1.0 to >32 .mu.g/mL), and CAP18-22 (MIC, 0.5 to 32 .mu.g/mL) had variable activities. LL37, mCRAMP, and rCRAMP were least active against the clin. isolates studied (MIC, 1.0 to >32 .mu.g/mL). Peptides had modest activities against *S. maltophilia* and *A. xylosoxidans* (MIC range, 1.0 to >32 .mu.g/mL), but none inhibited *B. cepacia*. However, CF sputum inhibited the activity of SMAP29 substantially. The effects of peptides on bacterial cell membranes and eukaryotic cells were examd. by SEM and by measuring transepithelial cell resistance, resp. SMAP29 caused the appearance of bacterial membrane blebs within 1 min, killed *P. aeruginosa* within 1 h, and caused a dose-dependent, reversible decrease in transepithelial resistance within 5 h. The tested cathelicidin-derived peptides represent a novel class of antimicrobial agents and warrant further development as prophylactic or therapeutic agents for CF lung disease.

REFERENCE 3: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

E1	1	LRRIGDE/SQEP
E2	1	LRRIGDELD/SQEP
E3	1 -->	LRRIIRKIIHIK/SQEP
E4	2	LRRIIRKIIHIKK/SQEP
E5	1	LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSAAFYFSNFLQQ TKGKPPSWN'AAA'PSPFLAFDPTKGL/SQEP
E6	1	LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSAAFYFSNFLQQ TKGKPPSWNVSPFLAFDPTKGL/SQEP
E7	1	LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSRGAYFFDFLQQ TKGKPPSWNVSPFLAFDP'AAA-AAA'GL/SQEP
E8	1	LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSRGAYFFDFLQQ TKGKPPSWNVSPFLAFDPTKGL/SQEP
E9	1	LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSRGAYFSDFLQQ TKGKPPSWNVSPFLAFDPT'AAA'GL/SQEP
E10	1	LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSRGAYFSDFLQQ TKGKPPSWNVSPFLAFDPTKGL/SQEP
E11	1	LRRILRGCAQRFIFEEVAPDQYAHTDASKMLRVTGIHALVGFSCDEVMSRGAYFSDFLQQ TKGNPPSLNVSPFLAFDPTKG'AAA'/SQEP

E12 3 LRRIR/SQEP

=> s e3;d sqide cbib abs;e rriirkiihiik/sqep

1 LRRIRKIIHIK/SQEP

292001 SQL=13

L7 1 (LRRIRKIIHIK)/SQEP

(LRRIRKIIHIK/SQEP AND SQL=13)

L7 ANSWER 1 OF 1 REGISTRY COPYRIGHT 2002 ACS

RN 326855-52-1 REGISTRY

CN L-Lysinamide, L-leucyl-L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-isoleucyl-L-isoleucyl-(9CI) (CA INDEX NAME)

OTHER NAMES:

CN 15: PN: WO0112668 SEQID: 15 claimed protein

CN Ovispirin OV 11

FS PROTEIN SEQUENCE; STEREOSEARCH

SQL 13

NTE modified

type	location	description
terminal mod.	Lys-13	C-terminal amide

SEQ 1 LRRIRKIIH IIK

=====

HITS AT: 1-13

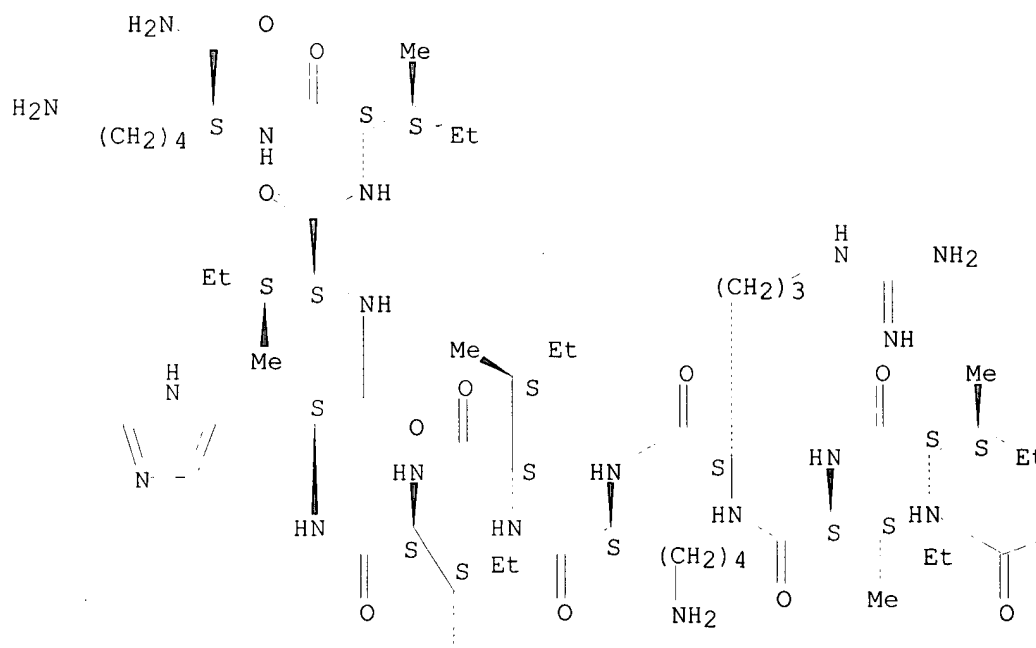
MF C78 H147 N27 O13

SR CA

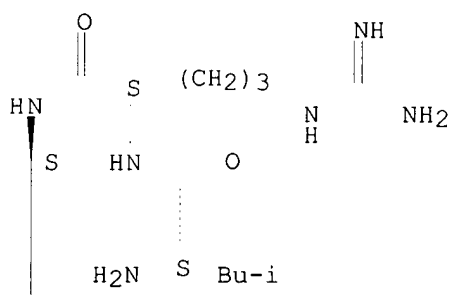
LC STN Files: CA, CAPLUS

Absolute stereochemistry.

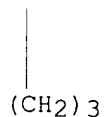
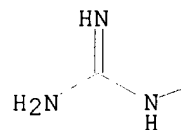
PAGE 1-A



Searched by: Mary Hale 308-4258 CM-1 1E01



Me



2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB SMAP29, an ovine cathelicidin, was systematically altered to create a

family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PA01.

REFERENCE 2: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

```
E1      1      RRIIPRGEARLPS/SQEP
E2      1      RRIIRKIIHII/SQEP
E3      1  --> RRIIRKIIHIK/SQEP
E4      1      RRIIRKIIHIKK/SQEP
E5      1      RRIIRPPNTTNQFAMTAGTSRPMARPFASRSSVLTWRKTLENPSMSEVPTGVFPVLVTP
          SWTTTGPFVPSSRSSLNRSVGWEPKHLKLRVPIWLSYGTS/SQEP
E6      11     RRIIRWKGMGELRREGAKRKEKILQKLITRKRMRDINFRQKEILSSATSSCQL/SQEP
E7      1      RRIIYIL/SQEP
E8      1      RRIKAITLK/SQEP
E9      1      RRIKDFLLTARRKDAKSVKIKKNKDNVFKFVRCSTRYLYTLV'AAA'TDKEKA'AAA'KLK
          QSLPPRFAQ/SQEP
E10     1      RRIKEEWGTSRWFLEFDIRKCFHTIDRHRFIPIFKEEIGDPKFFFYPTQKLFSAGRLVGGE
          KGPYSVPYSVLLSALSGNIYLHKLDQEIGRIRQKHEIPLVQIRIRSVLLRTGRRIDDOENS
          GEEASFNAPQDNRAIVIGRVKSIQRKATFHSLVSSWHTPPASTPRRRGDQKTTGVFPFPSS
          AKAFLNKPSSLLCAAFLE/SQEP
E11     1      RRIKEEWGTSRWFLEFDIRKCFHTIDRHRLIPIFKEEIDDPKFFFYPTQKVFSAGRLVGGE
          KGPYSVPHSVLLSALPGNIYLHKLDQEIGRIRQKYEIPIVQIRIRSVLLRTGRIDDOENSG
          EEASFNAPQDNRAIIVIGRVKSIQRKAFFHSLVSSWHTPPPTSTPRRRGDQKTPFVFPFSSA
          LAAFLNKPSSLLCAAFLEA/SQEP
E12     1      RRIKEIVKK/SQEP
```

```
=> s e3;d sqide cbib abs;e glrklrkrfrnkikeklkkig/sqep
      1 RRIIRKIIHIK/SQEP
      152033 SQL=12
L8     1 (RRIIRKIIHIK)/SQEP
      (RRIIRKIIHIK/SQEP AND SQL=12)
```

L8 ANSWER 1 OF 1 REGISTRY COPYRIGHT 2002 ACS

RN 326855-54-3 REGISTRY

CN L-Lysinamide, L-arginyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-arginyl-L-lysyl-L-isoleucyl-L-isoleucyl-L-histidyl-L-isoleucyl-L-isoleucyl- (9CI)
(CA INDEX NAME)

OTHER NAMES:

CN 17: PN: WO0112668 SEQID: 17 claimed protein

CN Ovispirin OV 13

FS PROTEIN SEQUENCE; STEREOSEARCH

SQL 12

NTE modified

type	-----	location	-----	description
terminal mod.	Lys-12	-		C-terminal amide

SEQ 1 RRIIRKIIHI IK

=====

HITS AT: 1-12

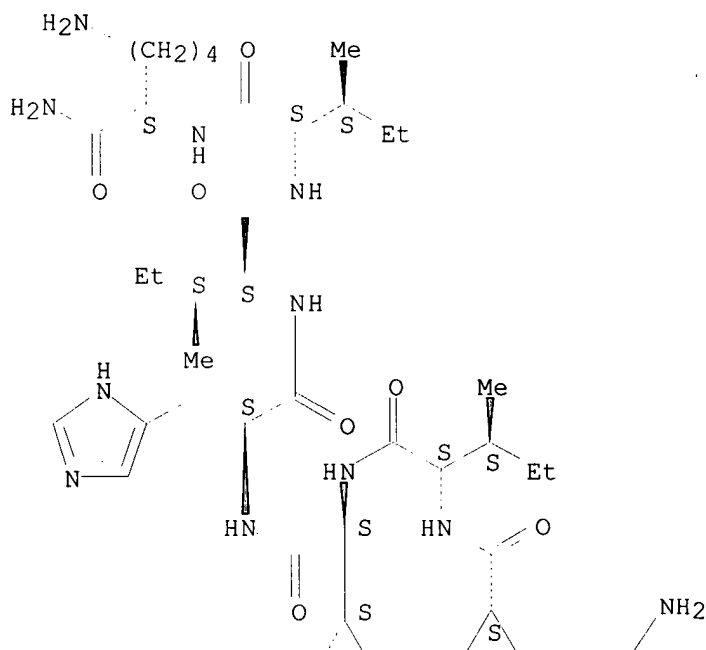
MF C72 H136 N26 O12

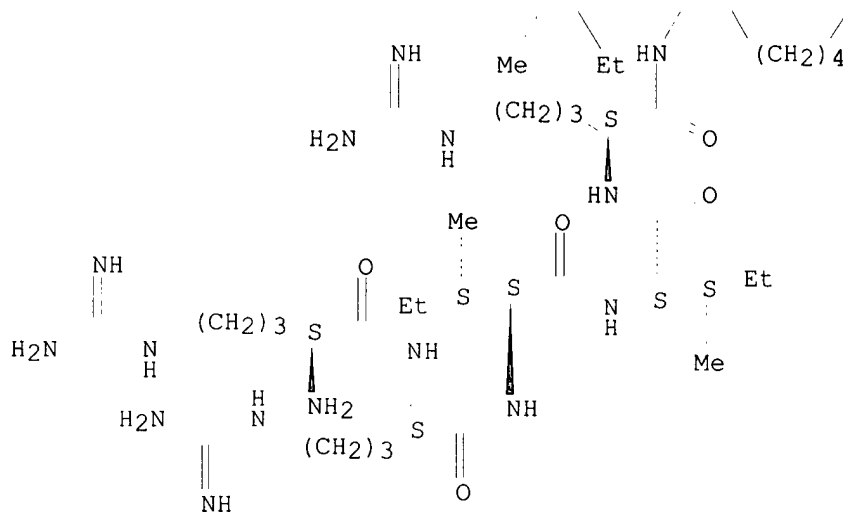
SR CA

LC STN Files: CA, CAPLUS

Absolute stereochemistry.

PAGE 1-A





2 REFERENCES IN FILE CA (1967 TO DATE)
2 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PAO1.

REFERENCE 2: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a

peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

```
E1      1      GLRKRLRKFRNKIKEK/SQEP
E2      2      GLRKRLRKFRNKIKEKLKKI/SQEP
E3      1  -->  GLRKRLRKFRNKIKEKLKKIG/SQEP
E4      1      GLRKRLRKFRNKIKEKLKKIGQ'AAA'IQ'AAA-AAA'LL/SQEP
E5      1      GLRKRLRKFRNKIKEKLKKIGQKIQ/SQEP
E6      1      GLRKRLRKFRNKIKEKLKKIGQKIQG/SQEP
E7      1      GLRKRLRKFRNKIKEKLKKIGQKIQGFVPKLAPRTDY/SQEP
E8      1      GLRKRLRKFRNKIKEKLKKIGQKIQGLL/SQEP
E9      2      GLRKRLRKFRNKIKEKLKKIGQKIQGLLPKLA/SQEP
E10     1      GLRKRLRKFRNKIKEKLKKIGQKIQGLLPKLAPRTDY/SQEP
E11     1      GLRKRLRKFRNKIKQKLKKI/SQEP
E12     1      GLRKRLRKPRNKPKEKLKKI/SQEP
```

```
=> s e3;d sqide cbib abs;e rkrlrkfrnkikeklkkigqki/sqep
      1 GLRKRLRKFRNKIKEKLKKIG/SQEP
70973 SQL=21
L9      1 (GLRKRLRKFRNKIKEKLKKIG)/SQEP
      (GLRKRLRKFRNKIKEKLKKIG/SQEP AND SQL=21)
```

```
L9  ANSWER 1 OF 1  REGISTRY  COPYRIGHT 2002 ACS
RN  276251-03-7  REGISTRY
CN  Glycine, glycyL-L-leucyl-L-arginyl-L-lysyl-L-arginyl-L-leucyl-L-arginyl-L-
    lysyl-L-phenylalanyl-L-arginyl-L-asparaginyL-L-lysyl-L-isoleucyl-L-lysyl-L-
    .alpha.-glutamyl-L-lysyl-L-leucyl-L-lysyl-L-lysyl-L-isoleucyl- (9CI)  (CA
    INDEX NAME)
```

OTHER NAMES:

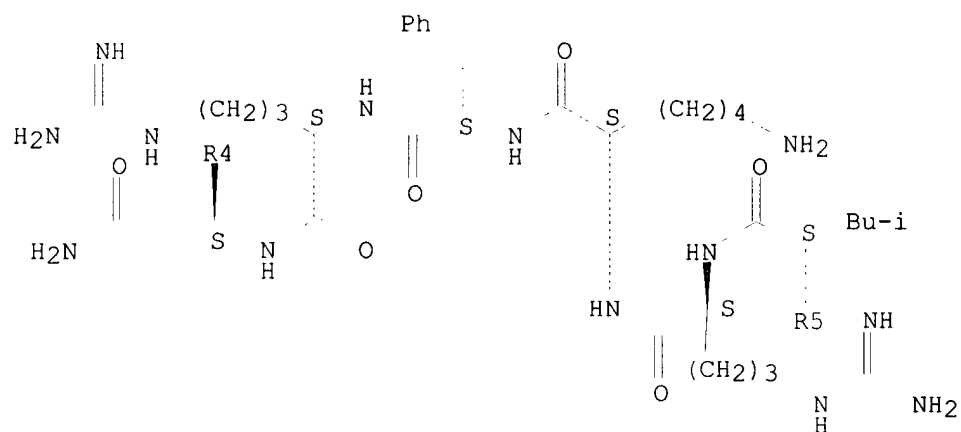
```
CN  19: PN: WO0112668 SEQID: 19 claimed protein
FS  PROTEIN SEQUENCE; STEREOSEARCH
SQL  21
```

```
SEQ      1 GLRKRLRKFR NKIKEKLKKI G
      =====
```

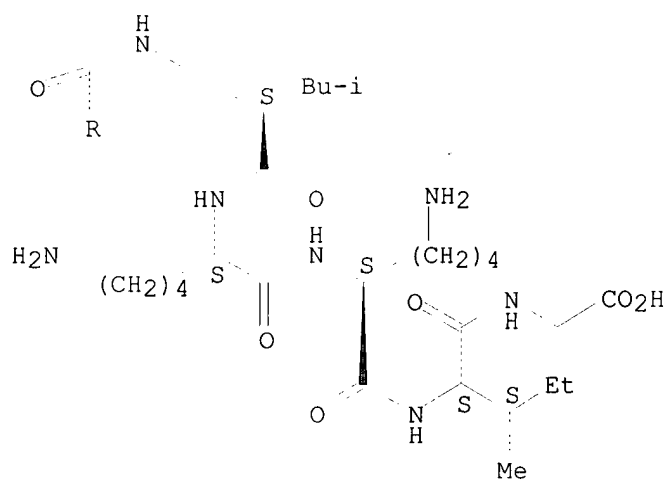
```
HITS AT:  1-21
MF  C118 H217 N41 O25
SR  CA
LC  STN Files:  CA, CAPLUS
```

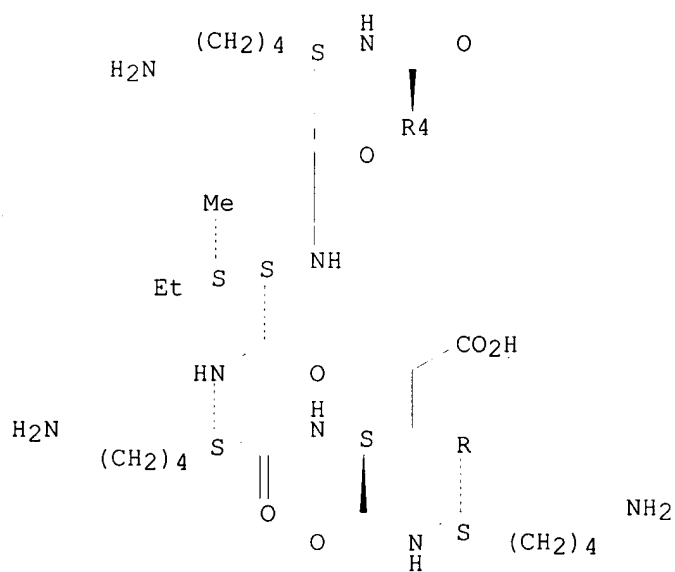
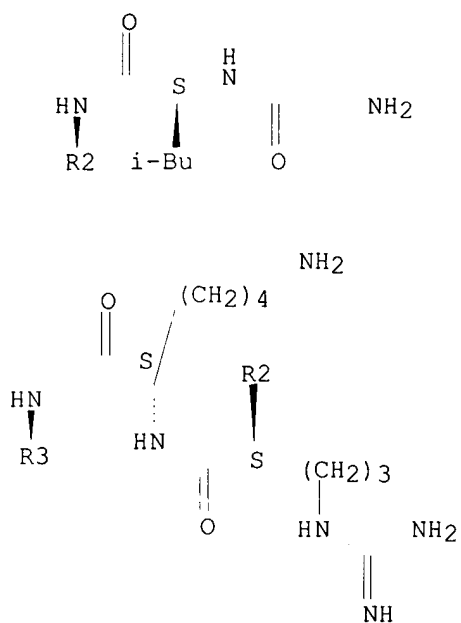
Absolute stereochemistry.

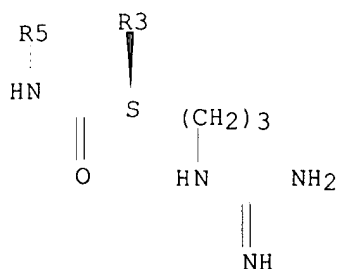
PAGE 1-A



PAGE 2-A







3 REFERENCES IN FILE CA (1967 TO DATE)

3 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 134:216813 The ovine cathelicidin SMAP29 kills ovine respiratory pathogens in vitro and in an ovine model of pulmonary infection. Brogden, K. A.; Kalfa, V. C.; Ackermann, M. R.; Palmquist, D. E.; McCray, P. B., Jr.; Tack, B. F. (Respiratory Diseases of Livestock Research Unit, USDA Agricultural Research Service, National Animal Disease Center, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(1), 331-334 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB Cathelicidins are antimicrobial peptides from sheep (SMAP29 and SMAP34), rabbits (CAP11 and CAP18), rodents (CRAMP), and humans (FALL39, LL37, and h/CAP18). In a broth microdilution assay against nine ovine pathogens, SMAP29, SMAP34, mouse CRAMP, CAP18, CAP1831, CAP1828, CAP1822, and CAP1821a were the most active, with MICs as low as 0.6 .mu.g/mL. Other cathelicidins were less active. In lambs with pneumonia, 0.5 mg of SMAP29 reduced the concn. of bacteria in both bronchoalveolar lavage fluid and consolidated pulmonary tissues. Hence, the antimicrobial activity of SMAP29 suggests that it has applications in the treatment of respiratory tract infections.

REFERENCE 2: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

REFERENCE 3: 133:57515 Bactericidal activity of mammalian

cathelicidin-derived peptides. Travis, Sue M.; Anderson, Norma N.; Forsyth, William R.; Espiritu, Cesar; Conway, Barbara D.; Greenberg, E. P.; McCray, Paul B., Jr.; Lehrer, Robert I.; Welsh, Michael J.; Tack, Brian F. (Department Internal Medicine, University of Iowa College of Medicine Iowa City, Iowa City, IA, 52242, USA). *Infection and Immunity*, 68(5), 2748-2755 (English) 2000. CODEN: INFIBR. ISSN: 0019-9567. Publisher: American Society for Microbiology.

AB Endogenous antimicrobial peptides of the cathelicidin family contribute to innate immunity. The emergence of widespread antibiotic resistance in many commonly encountered bacteria requires the search for new bactericidal agents with therapeutic potential. Solid-phase synthesis was employed to prep. linear antimicrobial peptides found in cathelicidins of five mammals: human (FALL39/LL37), rabbit (CAP18), mouse (mCRAMP), rat (rCRAMP), and sheep (SMAP29 and SMAP34). These peptides were tested at ionic strengths of 25 and 175 mM against *Pseudomonas aeruginosa*, *Escherichia coli*, *Staphylococcus aureus*, and methicillin-resistant *Staphylococcus aureus*. Each peptide manifested activity against *P. aeruginosa* irresp. of the NaCl concn. CAP18 and SMAP29 were the most effective peptides of the group against all test organisms under both low-and high-salt conditions. Select peptides of 15 to 21 residues, modeled on CAP18 (37 residues), retained activity against the gram-neg. bacteria and methicillin-sensitive *S. aureus*, although the bactericidal activity was reduced compared to that of the parent peptide. In accordance with the behavior of the parent mol., the truncated peptides adopted an .alpha.-helical structure in the presence of trifluoroethanol or lipopolysaccharide. The relation between the bactericidal activity and several physiochem. properties of the cathelicidins was examd. The activities of the full-length peptides correlated pos. with a predicted gradient of hydrophobicity along the peptide backbone and with net pos. charge; they correlated inversely with relative abundance of anionic residues. The salt-resistant, antimicrobial properties of CAP18 and SMAP29 suggest that these peptides or congeneric structures have potential for the treatment of bacterial infections in normal and immunocompromised persons and individuals with cystic fibrosis.

E1	1	RKRLQVQSLIAT/SQEP
E2	1	RKRLRKFRNKIKEKLKKIGQK/SQEP
E3	1 -->	RKRLRKFRNKIKEKLKKIGQKI/SQEP
E4	1	RKRLRLIHLHQT/SQEP
E5	1	RKRLRTKSLP/SQEP
E6	1	RKRMANICIATTATNQNQVILRDYVSGFPKESDLYIAAATVDLRVPPGSM AVLK NLYLSC DPHSRTRMGKPDPSPTSM AHAFTIGKPI SGFGVAKAIHSCHPNYKTG DLLWGRV GWE EY CVISPTPSSHFKIHHTDVPLSFYTGILGIPGLTAYVGFYEICSPKKGKTVFVSAASGAVG QLVGQFAKMVG CYVVG SASS/SQEP
E7	1	RKRMRNRRIAASKCRKRLER IAR/SQEP
E8	1	RKR NKAR/SQEP
E9	1	RKR NKARYSPLHNEDEAGDEDEL/SQEP
E10	1	RKR NQGGEAPSNDKAQGWAPAA ILYRAAGYSPASSALGGGGAAAEARKVARDASSRRRG 'AAA'NADWDGGGSSGTRRA/SQEP
E11	1	RKR NRIPLGCTICRKRKVKCDKLRPHCQQCTKTGVAHLCHYMEQ TWAE EAEKELLKD NEL KKLRER VKSLEKTL SK/SQEP
E12	1	RKR NRIPLSCTICRKRKVKCDKLRPHCQQCTKTGVAHLCHYMEQ TWAE EAEKELLKD NEL KKLRER VKSLEK/SQEP

=> s e3;d sqide cbib abs;e rglrrlgrkiahgvkkygptvlriiria/sqep
 1 RKRLRKFRNKIKEKLKKIGQKI/SQEP
 41188 SQL=22
 L10 1 (RKRLRKFRNKIKEKLKKIGQKI)/SQEP
 (RKRLRKFRNKIKEKLKKIGQKI/SQEP AND SQL=22)

L10 ANSWER 1 OF 1 REGISTRY COPYRIGHT 2002 ACS

RN 276251-01-5 REGISTRY

CN L-Isoleucine, L-arginyl-L-lysyl-L-arginyl-L-leucyl-L-arginyl-L-lysyl-L-phenylalanyl-L-arginyl-L-asparaginyl-L-lysyl-L-isoleucyl-L-lysyl-L-.alpha.-glutamyl-L-lysyl-L-leucyl-L-lysyl-L-lysyl-L-isoleucylglycyl-L-glutaminyl-L-lysyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 21: PN: WO0112668 SEQID: 21 claimed protein

FS PROTEIN SEQUENCE; STEREOSEARCH

SQL 22

SEQ 1 RKRLRKFRNK IKEKLKKIGQ KI

=====

HITS AT: 1-22

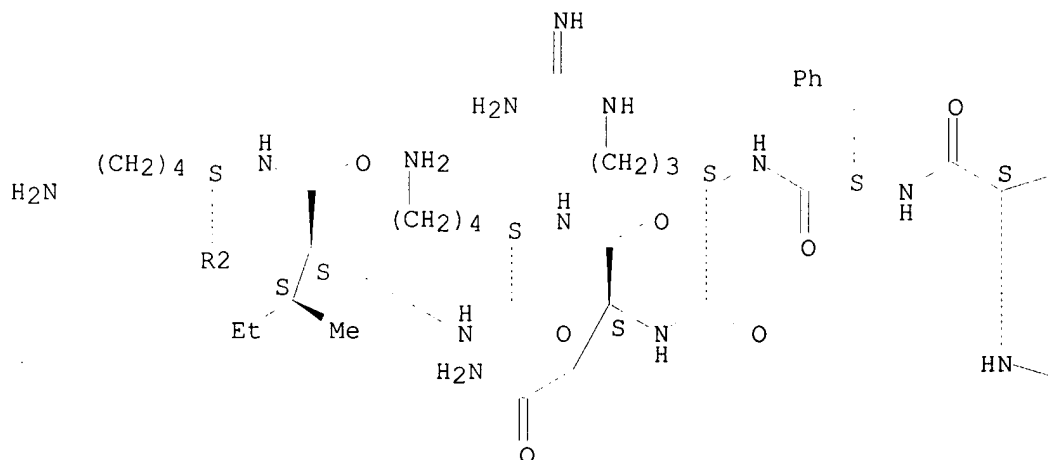
MF C127 H234 N44 O27

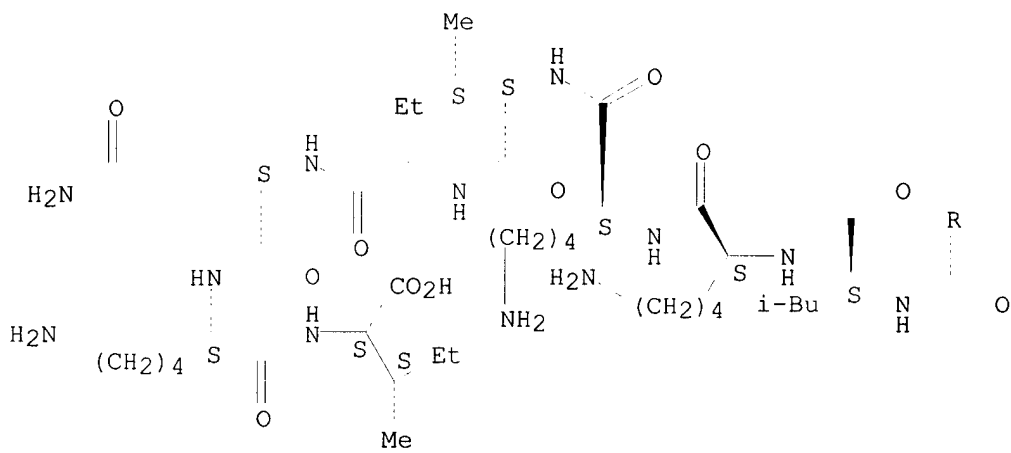
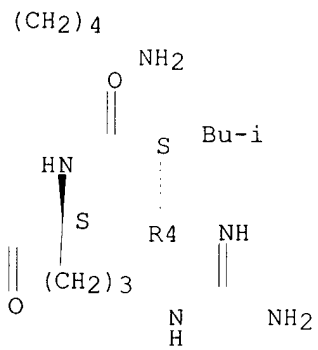
SR CA

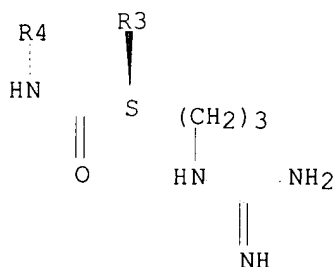
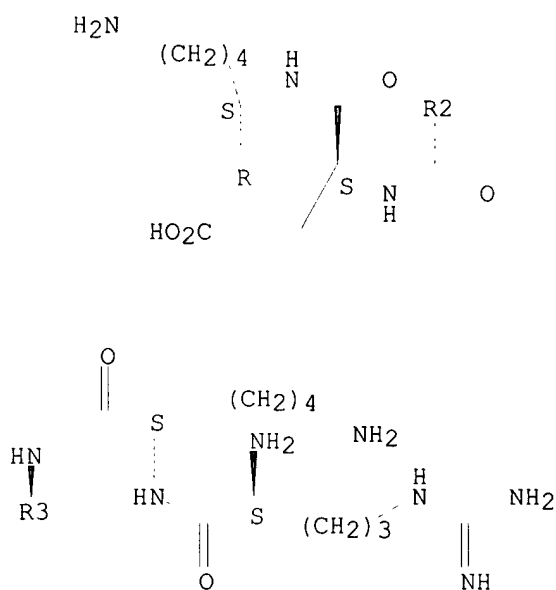
LC STN Files: CA, CAPLUS

Absolute stereochemistry.

PAGE 1-A







4 REFERENCES IN FILE CA (1967 TO DATE)
 4 REFERENCES IN FILE CAPLUS (1967 TO DATE)

- REFERENCE 1: 136:2776 Cathelicidin peptides inhibit multiply antibiotic-resistant pathogens from patients with cystic fibrosis. Saiman, Lisa; Tabibi, Setareh; Starner, Timothy D.; San Gabriel, Pablo; Winokur, Patricia L.; Jia, Hong Peng; McCray, Paul B., Jr.; Tack, Brian F. (Department of Pediatrics, Columbia University, New York, NY, 10032, USA). Antimicrobial Agents and Chemotherapy, 45(10), 2838-2844 (English) 2001. CODEN: AMACQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.
- AB Endogenous peptide antibiotics are under investigation as inhaled therapeutic agents for cystic fibrosis (CF) lung disease. The bactericidal activities of five cathelicidin peptides (LL37 [human], CAP18 [rabbit], mCRAMP [mouse], rCRAMP [rat], and SMAP29 [sheep]), three novel alpha-helical peptides derived from SMAP29 and termed ovispirins (OV-1, OV-2, and OV-3), and two derivs. of CAP18 were tested by broth microdilution assays. Their MICs were detd. for multiply antibiotic-resistant *Pseudomonas aeruginosa* (n = 24), *Burkholderia cepacia* (n = 5), *Achromobacter xylosoxidans* (n = 5), and *Stenotrophomonas maltophilia* (n = 5) strains isolated from CF patients. SMAP29 was most active and inhibited mucoid and nonmucoid *P. aeruginosa* strains (MIC, 0.06

to 8 .mu.g/mL). OV-1, OV-2, and OV-3 were nearly as active (MIC, 0.03 to 16 .mu.g/mL), but CAP18 (MIC, 1.0 to 32 .mu.g/mL), CAP18-18 (MIC, 1.0 to >32 .mu.g/mL), and CAP18-22 (MIC, 0.5 to 32 .mu.g/mL) had variable activities. LL37, mCRAMP, and rCRAMP were least active against the clin. isolates studied (MIC, 1.0 to >32 .mu.g/mL). Peptides had modest activities against *S. maltophilia* and *A. xylosoxidans* (MIC range, 1.0 to > 32 .mu.g/mL), but none inhibited *B. cepacia*. However, CF sputum inhibited the activity of SMAP29 substantially. The effects of peptides on bacterial cell membranes and eukaryotic cells were examd. by SEM and by measuring transepithelial cell resistance, resp. SMAP29 caused the appearance of bacterial membrane blebs within 1 min, killed *P. aeruginosa* within 1 h, and caused a dose-dependent, reversible decrease in transepithelial resistance within 5 h. The tested cathelicidin-derived peptides represent a novel class of antimicrobial agents and warrant further development as prophylactic or therapeutic agents for CF lung disease.

REFERENCE 2: 134:216813 The ovine cathelicidin SMAP29 kills ovine respiratory pathogens in vitro and in an ovine model of pulmonary infection. Brogden, K. A.; Kalfa, V. C.; Ackermann, M. R.; Palmquist, D. E.; McCray, P. B., Jr.; Tack, B. F. (Respiratory Diseases of Livestock Research Unit, USDA Agricultural Research Service, National Animal Disease Center, Ames, IA, 50010, USA). *Antimicrobial Agents and Chemotherapy*, 45(1), 331-334 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB Cathelicidins are antimicrobial peptides from sheep (SMAP29 and SMAP34), rabbits (CAP11 and CAP18), rodents (CRAMP), and humans (FALL39, LL37, and h/CAP18). In a broth microdilution assay against nine ovine pathogens, SMAP29, SMAP34, mouse CRAMP, CAP18, CAP1831, CAP1828, CAP1822, and CAP1821a were the most active, with MICs as low as 0.6 .mu.g/mL. Other cathelicidins were less active. In lambs with pneumonia, 0.5 mg of SMAP29 reduced the concn. of bacteria in both bronchoalveolar lavage fluid and consolidated pulmonary tissues. Hence, the antimicrobial activity of SMAP29 suggests that it has applications in the treatment of respiratory tract infections.

REFERENCE 3: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). *PCT Int. Appl. WO* 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

REFERENCE 4: 133:57515 Bactericidal activity of mammalian

cathelicidin-derived peptides. Travis, Sue M.; Anderson, Norma N.; Forsyth, William R.; Espiritu, Cesar; Conway, Barbara D.; Greenberg, E. P.; McCray, Paul B., Jr.; Lehrer, Robert I.; Welsh, Michael J.; Tack, Brian F. (Department Internal Medicine, University of Iowa College of Medicine Iowa City, Iowa City, IA, 52242, USA). Infection and Immunity, 68(5), 2748-2755 (English) 2000. CODEN: INFIBR. ISSN: 0019-9567. Publisher: American Society for Microbiology.

AB Endogenous antimicrobial peptides of the cathelicidin family contribute to innate immunity. The emergence of widespread antibiotic resistance in many commonly encountered bacteria requires the search for new bactericidal agents with therapeutic potential. Solid-phase synthesis was employed to prep. linear antimicrobial peptides found in cathelicidins of five mammals: human (FALL39/LL37), rabbit (CAP18), mouse (mCRAMP), rat (rCRAMP), and sheep (SMAP29 and SMAP34). These peptides were tested at ionic strengths of 25 and 175 mM against *Pseudomonas aeruginosa*, *Escherichia coli*, *Staphylococcus aureus*, and methicillin-resistant *Staphylococcus aureus*. Each peptide manifested activity against *P. aeruginosa* irresp. of the NaCl concn. CAP18 and SMAP29 were the most effective peptides of the group against all test organisms under both low- and high-salt conditions. Select peptides of 15 to 21 residues, modeled on CAP18 (37 residues), retained activity against the gram-neg. bacteria and methicillin-sensitive *S. aureus*, although the bactericidal activity was reduced compared to that of the parent peptide. In accordance with the behavior of the parent mol., the truncated peptides adopted an α -helical structure in the presence of trifluoroethanol or lipopolysaccharide. The relation between the bactericidal activity and several physiochem. properties of the cathelicidins was examd. The activities of the full-length peptides correlated pos. with a predicted gradient of hydrophobicity along the peptide backbone and with net pos. charge; they correlated inversely with relative abundance of anionic residues. The salt-resistant, antimicrobial properties of CAP18 and SMAP29 suggest that these peptides or congeneric structures have potential for the treatment of bacterial infections in normal and immunocompromised persons and individuals with cystic fibrosis.

```

E1      1      RGLRRLGRKIAHG VKKYGATVLR IIRIA/SQEP
E2      1      RGLRRLGRKIAHG VKKYGPTVKRIKRKA/SQEP
E3      1  -->  RGLRRLGRKIAHG VKKYGPTVLR IIRIA/SQEP
E4      1      RGLRRLGRKIAHG VKKYGPTVLR IIRIAG/SQEP
E5      1      RGLRSASGSCGSRGRGGLSTGSRPCLDGLLRRSGLLAGSLLVGAEQSHDPSFYVSSSRAR
          WLVPAGATLSGLRRLYGYCCCRCAVPYPNDQATATLTTSRSPQAPKGKDSTPPMAGIFA
          RLSTVFRSKANKALDKVEDPRETLDYSYQRQLDMLQKVRRGVADVATSRKRVELQMNQLQ
          TDVDKMQNAASRALEQGRED/SQEP
E6      1      RGLRSGTYGD/SQEP
E7      1      RGLRSRWLPQSRRHCRARENQDSSWAAARRTASTPTTPARHQRRMPVTRSSRSERLDDRQ
          PER/SQEP
E8      1      RGLRWIFRRCFR/SQEP
E9      1      RGLS'AAA'HGPAGVRGSAGAAFFITYEYVKWFLHADSSSYLTMPKHMLAASAGEVVAC
          LIRVPSEVVKQRAQVSASTRTFQIFSNILYEEGIQGLYRGYKSTVLREIPFSLVQFPLWE
          SLKALWSW'AAA'QDHVVDWSQSAVCGAFAGGFAAAV/SQEP
E10     1      RGLSAIRERL/SQEP
E11     1      RGLSASMTGARCDAAALTLWLGLLLMNQVRIRYGLLIGYSQLKVGRLLRPTEKVMMSGNGN
          RCKLCTT/SQEP
E12     1      RGLSDDASHAETPSPLTPSRVDSFSDGVERTRRSSPRVEEHQTSSREEKAATERVPKLSR
          LPSLRAPLRSTD RRASPPRRLSQLLRCTTSRFASKGTAYPDEEWGHRVRAQRTEETVSS
          LTTKRLLTRSPNSQTAFPR/SQEP

```

```

=> s e3;d sqide cbib abs;dis his
      1 RGLRRLGRKIAHG VKKYGPTVLR IIRIA/SQEP
      22834 SQL=28

```

L11 1 (RGLRRLGRKIAHGVKKYGPTVLRIIRIA)/SQEP
(RGLRRLGRKIAHGVKKYGPTVLRIIRIA/SQEP AND SQL=28)

L11 ANSWER 1 OF 1 REGISTRY COPYRIGHT 2002 ACS

RN 255817-49-3 REGISTRY

CN L-Alaninamide, L-arginylglycyl-L-leucyl-L-arginyl-L-arginyl-L-leucylglycyl-L-arginyl-L-lysyl-L-isoleucyl-L-alanyl-L-histidylglycyl-L-valyl-L-lysyl-L-lysyl-L-tyrosylglycyl-L-prolyl-L-threonyl-L-valyl-L-leucyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-arginyl-L-isoleucyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 27: PN: WO0112668 SEQID: 27 claimed protein

CN SMAP 28

FS PROTEIN SEQUENCE; STEREOSEARCH

SQL 28

NTE modified

type	location	description
terminal mod.	Ala-28	C-terminal amide

SEQ 1 RGLRRLGRKI AHGVKKYGPT VLRIIRIA

=====

HITS AT: 1-28

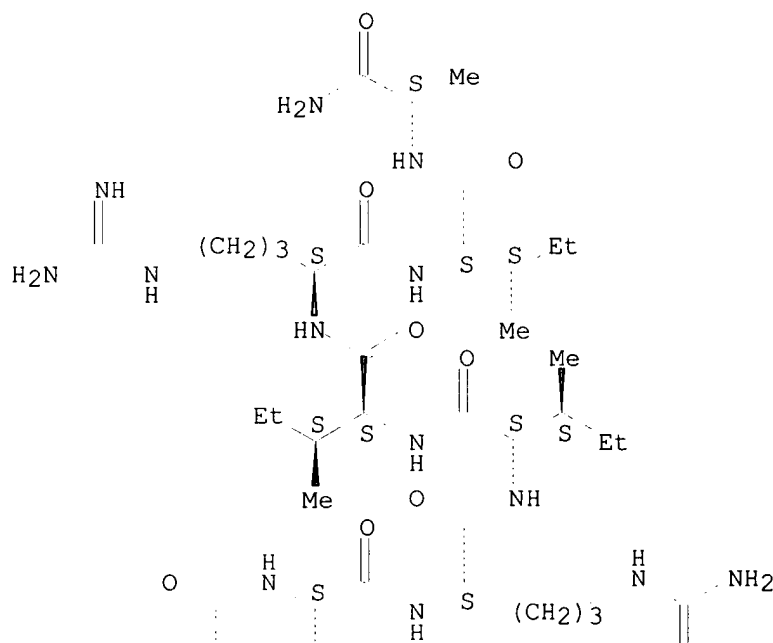
MF C144 H258 N52 O30

SR CA

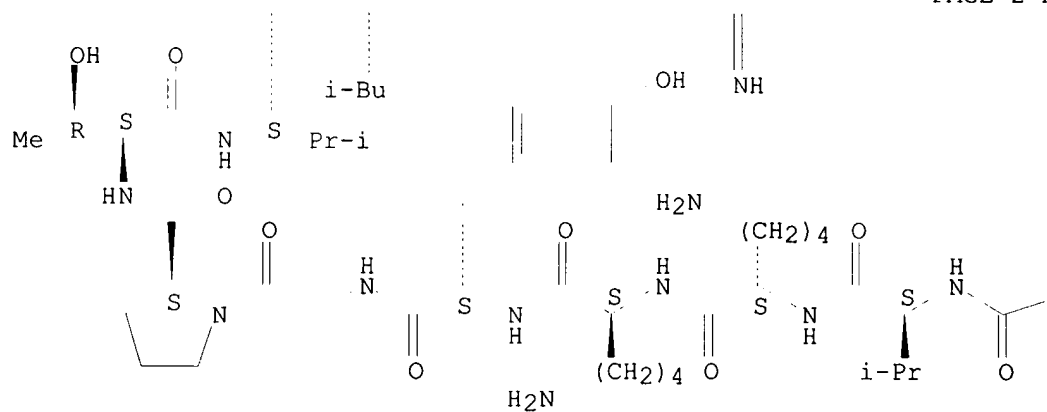
LC STN Files: CA, CAPLUS

Absolute stereochemistry.

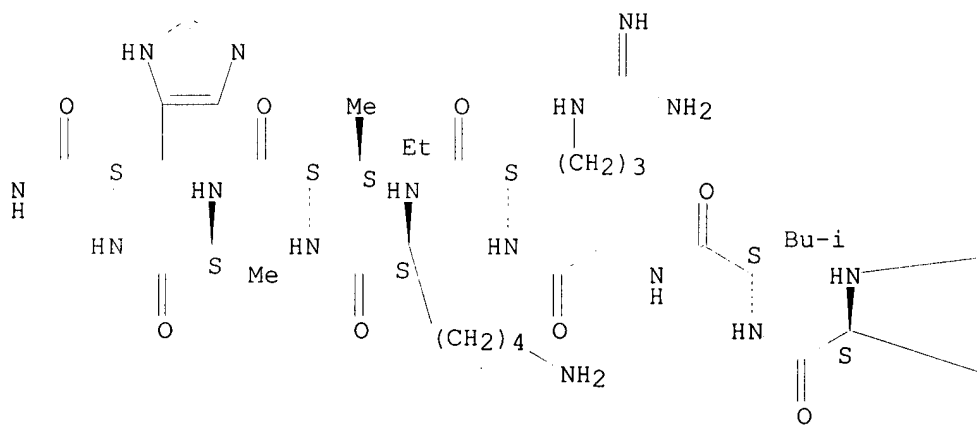
PAGE 1-A

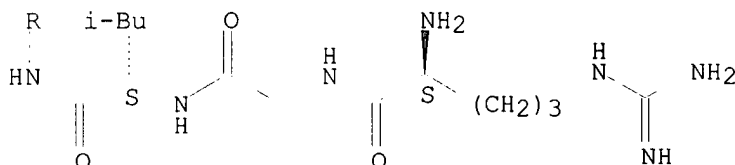
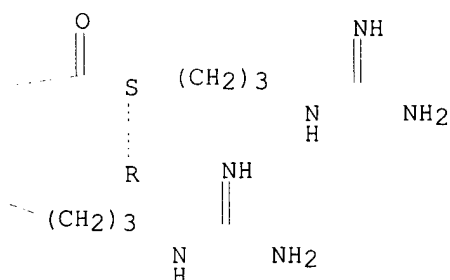


PAGE 2-A



PAGE 2-B





3 REFERENCES IN FILE CA (1967 TO DATE)

3 REFERENCES IN FILE CAPLUS (1967 TO DATE)

REFERENCE 1: 136:82518 Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Kalfa, V. C.; Jia, H. P.; Kunkle, R. A.; McCray, P. B., Jr.; Tack, B. F.; Brogden, K. A. (Respiratory Diseases of Livestock Research Unit, National Animal Disease Center, USDA Agricultural Research Service, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3256-3261 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB SMAP29, an ovine cathelicidin, was systematically altered to create a family of 23 related peptides for MIC and min. bactericidal concn. detns. SMAP28, SMAP29, and a deriv. of SMAP29 called ovispirin were all antimicrobial. However, many congeners of SMAP29 and ovispirin were not as active as the parent mols. With immunoelectron microscopy, SMAP29 was seen on membranes and within the cytoplasm of *Pseudomonas aeruginosa* PA01.

REFERENCE 2: 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO,

NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2.
APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

REFERENCE 3: 132:106875 SMAP-29: a potent antibacterial and antifungal peptide from sheep leukocytes. Skerlavaj, B.; Benincasa, M.; Risso, A.; Zanetti, M.; Gennaro, R. (Dipartimento di Scienze e Tecnologie Biomediche, Universita di Udine, Udine, 33100, Italy). FEBS Letters, 463(1,2), 58-62 (English) 1999. CODEN: FEBLAL. ISSN: 0014-5793. Publisher: Elsevier Science B.V..

AB SMAP-29 is a cathelicidin-derived peptide deduced from sheep m... The C-terminally amidated form of this peptide was chem. syntl shown to exert a potent antimicrobial activity. Antibiotic-re clin. isolates highly susceptible to this peptide include MRSA isolates, that are a major worldwide problem, and mucoid Pseud aeruginosa assocd. with chronic respiratory inflammation in CF. In addn., SMAP-29 is also active against fungi, including Crypt neoformans isolated from immunocompromised patients. SMAP-29 c significant morphol. alterations of the bacterial surfaces, as SEM, and is also hemolytic against human, but not sheep erythro potent antimicrobial activity suggests that this peptide is an candidate as a lead compd. for the development of novel antiinf agents.

(FILE 'HOME' ENTERED AT 08:51:59 ON 12 JUL 2002)

FILE 'REGISTRY' ENTERED AT 09:00:30 ON 12 JUL 2002

	E KNLRRRIIRKIIHIKKYG/SQEP
L1	2 S E3
L2	2 S E3
	E KNIRRIIRKIIHIKKYG/SQEP
L3	2 S E3
	E NLRRRIIRKIIHIKKY/SQEP
L4	2 S E3
	E LRRIIRKIIHIKK/SQEP
L5	1 S E2
	E IRRRIIRKIIHIKK/SQEP
L6	2 S E3
	E LRRIIRKIIHIK/SQEP
L7	1 S E3
	E RRIIRKIIHIK/SQEP
L8	1 S E3
	E GLRKRLRKFRNKIKEKLKKIG/SQEP
L9	1 S E3
	E RKRLRKFRNKIKEKLKKIGQKI/SQEP
L10	1 S E3

Searched by: Mary Hale 308-4258 CM-1 1E01

E RGLRRLGRKIAHGVKKYGPTVLRIRIA/SQEP

L11 1 S E3

=> fil medl,hcap,biosis,embase,jicst;s (l1 or l2 or l3 or l4 or l5 or l6 or l7 or l8 or l9 or l10 or l11) and (antimicrobial or antibacterial)

COST IN U.S. DOLLARS	SINCE FILE ENTRY	TOTAL SESSION
FULL ESTIMATED COST	183.32	186.26

DISCOUNT AMOUNTS (FOR QUALIFYING ACCOUNTS)	SINCE FILE ENTRY	TOTAL SESSION
CA SUBSCRIBER PRICE	-8.26	-8.26

FILE 'MEDLINE' ENTERED AT 09:09:49 ON 12 JUL 2002

FILE 'HCAPLUS' ENTERED AT 09:09:49 ON 12 JUL 2002

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'BIOSIS' ENTERED AT 09:09:49 ON 12 JUL 2002

COPYRIGHT (C) 2002 BIOLOGICAL ABSTRACTS INC. (R)

FILE 'EMBASE' ENTERED AT 09:09:49 ON 12 JUL 2002

COPYRIGHT (C) 2002 Elsevier Science B.V. All rights reserved.

FILE 'JICST-EPLUS' ENTERED AT 09:09:49 ON 12 JUL 2002

COPYRIGHT (C) 2002 Japan Science and Technology Corporation (JST)

L12	0 FILE MEDLINE
L13	8 FILE HCAPLUS
L14	0 FILE BIOSIS
L15	0 FILE EMBASE
L16	0 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L17 8 (L1 OR L2 OR L3 OR L4 OR L5 OR L6 OR L7 OR L8 OR L9 OR L10 OR L11) AND (ANTIMICROBIAL OR ANTIBACTERIAL)

=> s l17 and (smap 209 or rcap 18)

L18	0 FILE MEDLINE
L19	1 FILE HCAPLUS
L20	0 FILE BIOSIS
L21	0 FILE EMBASE
L22	0 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L23 1 L17 AND (SMAP 209 OR RCAP 18)

=> d cbib abs

L23 ANSWER 1 OF 1 HCAPLUS COPYRIGHT 2002 ACS

2001:137250 Document No. 134:188181 Cathelicidin-derived peptides with broad spectrum **antimicrobial** activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,

Searched by: Mary Hale 308-4258 CM-1 1E01

TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM;
RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB,
GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English).
CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US
1999-PV149886 19990818.

AB The invention relates to the use of **antimicrobial** peptides in
the inhibition of microbial growth and proliferation.
Antimicrobial truncated peptides are disclosed which are based on
SMAP 29 and **RCAP 18**, but which contain a lesser no. of
amino acid residues yet still retain bactericidal activity. In addn.,
synthetic peptides based upon the SMAP 29 protein are disclosed which have
fewer amino acid residues and include substitutions yet retain substantial
activity. The invention also relates to a method of inhibiting microbial
growth by administering an effective amt. of a peptide in accordance with
the invention, or by combining the peptides with other
antimicrobial agents or antibiotics.

=> s ((microbial growth and prolifer?) or (antimicrobial or microbial)(w)peptide?)
and (smap 29 or rcap 18)

L24 8 FILE MEDLINE
L25 12 FILE HCAPLUS
L26 5 FILE BIOSIS
L27 6 FILE EMBASE
L28 0 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L29 31 ((MICROBIAL GROWTH AND PROLIFER?) OR (ANTIMICROBIAL OR MICROBIAL
) (W) PEPTIDE?) AND (SMAP 29 OR RCAP 18)

=> s l29 and (antimicrobial agent? or antibiotic?)

L30 1 FILE MEDLINE
L31 7 FILE HCAPLUS
L32 1 FILE BIOSIS
L33 2 FILE EMBASE
L34 0 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L35 11 L29 AND (ANTIMICROBIAL AGENT? OR ANTIBIOTIC?)

=> s l35 and (pharmaceutical or alpha helical peptide!)

L36 0 FILE MEDLINE
L37 0 FILE HCAPLUS
L38 0 FILE BIOSIS
L39 0 FILE EMBASE
L40 0 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L41 0 L35 AND (PHARMACEUTICAL OR ALPHA HELICAL PEPTIDE!)

=> dup rem l35

PROCESSING COMPLETED FOR L35

L42 8 DUP REM L35 (3 DUPLICATES REMOVED)

=> d cbib abs 1-8;s tack, b?/au,in;s mccray, p?/au,in;s welsh, m?/au,in;s lehrer,
r?/au,in

L42 ANSWER 1 OF 8 HCAPLUS COPYRIGHT 2002 ACS

2002:177892 Document No. 136:275091 **SMAP-29** has two

LPS-binding sites and a central hinge. Tack, Brian F.; Sawai, Monali V.;
Kearney, William R.; Robertson, Andrew D.; Sherman, Mark A.; Wang, Wei;
Hong, Teresa; Boo, Lee Ming; Wu, Huiyuan; Waring, Alan J.; Lehrer, Robert

Searched by: Mary Hale 308-4258 CM-1 1E01

I. (Department of Microbiology, University of Iowa, IA, USA). European Journal of Biochemistry, 269(4), 1181-1189 (English) 2002. CODEN: EJBCAI. ISSN: 0014-2956. Publisher: Blackwell Publishing Ltd..

AB The CD spectra of **SMAP-29**, an **antimicrobial peptide** from sheep, showed disordered structure in aq. buffers, and significant helicity in membrane-like environments, including SDS micelles, lipopolysaccharide (LPS) dispersions, and trifluoroethanol buffer systems. A structure detd. by NMR in 40% perdeuterated trifluoroethanol indicated that residues 8-17 were helical, residues 18-19 formed a hinge, and residues 20-28 formed an ordered, hydrophobic segment. **SMAP-29** was flexible in 40% trifluoroethanol, forming two sets of conformers that differed in the relative orientation of the N-terminal domain. We used a chromogenic Limulus assay to det. the EC50 of the peptide (the concn. that bound 50% of the added LPS). Studies with full-length and truncated **SMAP-29** mols. revealed that each end of the holopeptide contained an LPS-binding domain. The higher affinity LPS-binding domain was situated in the flexible N-terminal portion. LPS binding to full-length **SMAP-29** showed pos. cooperativity, so the EC50 of the peptide (2.6 .mu.M) was considerably lower than that of the individual LPS-binding domains. LPS-binding studies with a mixt. of truncated peptides revealed that this cooperativity was primarily intramol. (i.e. involving the N- and C-terminal LPS-binding sites of the same peptide mol.). CAP-18[106-142], an antimicrobial cathelicidin peptide of rabbits, resembled **SMAP-29** in that it contained N- and C-terminal LPS-binding domains, had an EC50 of 2.5 .mu.M, and bound LPS with pos. cooperativity. We conclude that the presence of multiple binding sites that function cooperatively allow peptides such as **SMAP-29** and CAP-18 to bind LPS with high affinity.

L42 ANSWER 2 OF 8 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.

2002172163 EMBASE Cathelicidin peptides as candidates for a novel class of antimicrobials. Zanetti M.; Gennaro R.; Skerlavaj B.; Tomasinsig L.; Circo R.. M. Zanetti, Dipt. di Scienze Technologie Biomed., Universita di Udine, P.le Kolbe 4, 33100 Udine, Italy. zanetti@icgeb.trieste.it. Current Pharmaceutical Design 8/9 (779-793) 2002. Refs: 135.

ISSN: 1381-6128. CODEN: CPDEFP. Pub. Country: Netherlands. Language: English. Summary Language: English.

AB Cathelicidin peptides are a numerous group of mammalian cationic **antimicrobial peptides**. Despite a common evolutionary origin of their genes, peptides display a remarkable variety of sizes, sequences and structures. Their spectra of antimicrobial activity are varied and cover a range of organisms that includes bacteria, fungi and enveloped viruses. In addition, they bind to and neutralize the effects of endotoxin. These features make this family of peptides good candidates in view of a therapeutic use. The most promising ones are currently under evaluation as leads for the development of novel anti-infectives, and synthetic variants are in an advanced stage of development for specific clinical applications. This review focuses on recent studies on the structure and in vitro and in vivo biological activities of these peptides.

L42 ANSWER 3 OF 8 HCAPLUS COPYRIGHT 2002 ACS

2001:137250 Document No. 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. Tack, Brian E.; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,

MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of **antimicrobial peptides** in the inhibition of **microbial growth** and **proliferation**. Antimicrobial truncated peptides are disclosed which are based on **SMAP 29** and **RCAP 18**, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the **SMAP 29** protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting **microbial growth** by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other **antimicrobial agents** or **antibiotics**.

L42 ANSWER 4 OF 8 HCAPLUS COPYRIGHT 2002 ACS

2001:799056 Document No. 136:68577 Susceptibilities of oral bacteria and yeast to mammalian cathelicidins. Guthmiller, Janet M.; Vargas, Kaaren G.; Srikantha, Rupasree; Schomberg, Lori L.; Weistroffer, Paula L.; McCray, Paul B., Jr.; Tack, Brian F. (Department of Periodontics and Dows Institute for Dental Research, College of Medicine, University of Iowa, Iowa City, IA, 52242, USA). Antimicrobial Agents and Chemotherapy, 45(11), 3216-3219 (English) 2001. CODEN: AMACCQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB The effects of cathelicidins against oral bacteria and clin. important oral yeasts are not known. We tested the susceptibilities of Actinobacillus actinomycetemcomitans, Fusobacterium nucleatum, Porphyromonas gingivalis, Streptococcus sanguis, Candida krusei, Candida tropicalis and Candida albicans to the following cathelicidins: FALL39, SMAP29, and CAP18. SMAP29 and CAP18 were antimicrobial, whereas FALL39 did not exhibit antimicrobial activity. Future studies are needed to det. the potential use of these **antimicrobial peptides** in prevention and treatment of oral infections.

L42 ANSWER 5 OF 8

MEDLINE

DUPLICATE 1

2001420242 Document Number: 21361089. PubMed ID: 11467858.

Structure-activity analysis of **SMAP-29**, a sheep leukocytes-derived **antimicrobial peptide**. Shin S Y; Park E J; Yang S T; Jung H J; Eom S H; Song W K; Kim Y; Hahm K S; Kim J I. (Department of Life Science, Kwangju Institute of Science and Technology, Kwangju, 500-712, Korea.) BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, (2001 Jul 27) 285 (4) 1046-51. Journal code: 0372516. ISSN: 0006-291X. Pub. country: United States. Language: English.

AB **SAMP-29** is a cathelicidin-derived **antimicrobial peptide** deduced from sheep myeloid mRNA. To elucidate the structural-activity relationship of **SMAP-29**, several analogues were synthesized and their **antibiotic** activity was investigated. Compared to parental **SMAP-29**, **SMAP-29(1-17)** and **[K(22,25,27)]-SMAP-29** retained relatively effective antimicrobial activity (MIC: 1.0-8.0 microM), but resulted in a complete loss of hemolytic activity. Pro-19 --> Ala substitution (**[A19]-SMAP-29**) in **SMAP-29** induced a significant reduction in antibacterial activity. These results suggested that the N-terminal amphipathic alpha-helical region and the C-terminal hydrophobic region of **SMAP-29** are responsible for antimicrobial activity and hemolytic activity, respectively, and the central Pro-19 in **SMAP-29** plays

a critical role in showing improved antibacterial activity. In particular, [K(2,7,13)]-**SMAP-29**(1-17) showed potent antimicrobial activity under high salt conditions without hemolytic activity. Thus, this short peptide could serve as an attractive candidate for the development of therapeutic antimicrobial drugs. Structural analysis by circular dichroism suggested that **SMAP-29** seems to adopt a helix-bend/turn-extended random conformation.
Copyright 2001 Academic Press.

L42 ANSWER 6 OF 8 HCAPLUS COPYRIGHT 2002 ACS

2001:15541 Document No. 134:216813 The ovine cathelicidin SMAP29 kills ovine respiratory pathogens in vitro and in an ovine model of pulmonary infection. Brogden, K. A.; Kalfa, V. C.; Ackermann, M. R.; Palmquist, D. E.; McCray, P. B., Jr.; Tack, B. F. (Respiratory Diseases of Livestock Research Unit, USDA Agricultural Research Service, National Animal Disease Center, Ames, IA, 50010, USA). Antimicrobial Agents and Chemotherapy, 45(1), 331-334 (English) 2001. CODEN: AMACQ. ISSN: 0066-4804. Publisher: American Society for Microbiology.

AB Cathelicidins are **antimicrobial peptides** from sheep (SMAP29 and SMAP34), rabbits (CAP11 and CAP18), rodents (CRAMP), and humans (FALL39, LL37, and h/CAP18). In a broth microdilution assay against nine ovine pathogens, SMAP29, SMAP34, mouse CRAMP, CAP18, CAP1831, CAP1828, CAP1822, and CAP1821a were the most active, with MICs as low as 0.6 .mu.g/mL. Other cathelicidins were less active. In lambs with pneumonia, 0.5 mg of SMAP29 reduced the concn. of bacteria in both bronchoalveolar lavage fluid and consolidated pulmonary tissues. Hence, the antimicrobial activity of SMAP29 suggests that it has applications in the treatment of respiratory tract infections.

L42 ANSWER 7 OF 8 HCAPLUS COPYRIGHT 2002 ACS

2000:282566 Document No. 133:57515 Bactericidal activity of mammalian cathelicidin-derived peptides. Travis, Sue M.; Anderson, Norma N.; Forsyth, William R.; Espiritu, Cesar; Conway, Barbara D.; Greenberg, E. P.; McCray, Paul B., Jr.; Lehrer, Robert I.; Welsh, Michael J.; Tack, Brian F. (Department Internal Medicine, University of Iowa College of Medicine Iowa City, Iowa City, IA, 52242, USA). Infection and Immunity, 68(5), 2748-2755 (English) 2000. CODEN: INFIBR. ISSN: 0019-9567. Publisher: American Society for Microbiology.

AB Endogenous **antimicrobial peptides** of the cathelicidin family contribute to innate immunity. The emergence of widespread **antibiotic** resistance in many commonly encountered bacteria requires the search for new bactericidal agents with therapeutic potential. Solid-phase synthesis was employed to prep. linear **antimicrobial peptides** found in cathelicidins of five mammals: human (FALL39/LL37), rabbit (CAP18), mouse (mCRAMP), rat (rCRAMP), and sheep (SMAP29 and SMAP34). These peptides were tested at ionic strengths of 25 and 175 mM against *Pseudomonas aeruginosa*, *Escherichia coli*, *Staphylococcus aureus*, and methicillin-resistant *Staphylococcus aureus*. Each peptide manifested activity against *P. aeruginosa* irresp. of the NaCl concn. CAP18 and SMAP29 were the most effective peptides of the group against all test organisms under both low-and high-salt conditions. Select peptides of 15 to 21 residues, modeled on CAP18 (37 residues), retained activity against the gram-neg. bacteria and methicillin-sensitive *S. aureus*, although the bactericidal activity was reduced compared to that of the parent peptide. In accordance with the behavior of the parent mol., the truncated peptides adopted an .alpha.-helical structure in the presence of trifluoroethanol or lipopolysaccharide. The relation between the bactericidal activity and several physiochem. properties of the cathelicidins was examd. The activities of the full-length peptides correlated pos. with a predicted gradient of hydrophobicity along the peptide backbone and with net pos. charge; they correlated inversely with relative abundance of anionic

residues. The salt-resistant, antimicrobial properties of CAP18 and SMAP29 suggest that these peptides or congeneric structures have potential for the treatment of bacterial infections in normal and immunocompromised persons and individuals with cystic fibrosis.

L42 ANSWER 8 OF 8 HCAPLUS COPYRIGHT 2002 ACS

1998:597848 Document No. 130:2113 Biological activity of a novel, cathelicidin-derived **antimicrobial peptide** from sheep. Merluzzi, Laura; Scocchi, Marco; Zanetti, Margherita; Bagella, Luigi; Gennaro, Renato (Department of Biomedical Science and Technology, University of Udine, Udine, I-33100, Italy). Peptides 1996, Proceedings of the European Peptide Symposium, 24th, Edinburgh, Sept. 8-13, 1996, Meeting Date 1996, 639-640. Editor(s): Ramage, Robert; Epton, Roger. Mayflower Scientific: Kingswinford, UK. (English) 1998. CODEN: 66RCA5.

AB **SMAP-29** was identified in sheep by RT-PCR-based search for the cDNA of cathelicidin homologs. Its structure suggested the **SMAP-29** may possess antimicrobial activity. In vitro tests showed that **SMAP-29** is a highly potent membrane-active peptide. Its broad spectrum of activity, which includes fungi and methicillin-resistant Staphylococcus aureus strains, makes **SMAP-29** a good candidate as a drug lead for the development of novel **antibiotic** compds.

'IN' IS NOT A VALID FIELD CODE

L43 94 FILE MEDLINE

L44 90 FILE HCAPLUS

L45 139 FILE BIOSIS

'IN' IS NOT A VALID FIELD CODE

L46 78 FILE EMBASE

L47 0 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L48 401 TACK, B?/AU, IN

'IN' IS NOT A VALID FIELD CODE

L49 80 FILE MEDLINE

L50 72 FILE HCAPLUS

L51 82 FILE BIOSIS

'IN' IS NOT A VALID FIELD CODE

L52 44 FILE EMBASE

L53 0 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L54 278 MCCRAY, P?/AU, IN

'IN' IS NOT A VALID FIELD CODE

L55 533 FILE MEDLINE

L56 467 FILE HCAPLUS

L57 847 FILE BIOSIS

'IN' IS NOT A VALID FIELD CODE

L58 479 FILE EMBASE

L59 1 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L60 2327 WELSH, M?/AU, IN

'IN' IS NOT A VALID FIELD CODE

L61 211 FILE MEDLINE
L62 246 FILE HCAPLUS
L63 319 FILE BIOSIS
'IN' IS NOT A VALID FIELD CODE
L64 170 FILE EMBASE
L65 0 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L66 946 LEHRER, R?/AU, IN

=> s 148 and 154 and 160 and 166

L67 1 FILE MEDLINE
L68 2 FILE HCAPLUS
L69 1 FILE BIOSIS
L70 1 FILE EMBASE
L71 0 FILE JICST-EPLUS

TOTAL FOR ALL FILES

L72 5 L48 AND L54 AND L60 AND L66

=> dup rem 172

PROCESSING COMPLETED FOR L72

L73 2 DUP REM L72 (3 DUPLICATES REMOVED)

=> d cbib abs 1-2

L73 ANSWER 1 OF 2 HCAPLUS COPYRIGHT 2002 ACS

2001:137250 Document No. 134:188181 Cathelicidin-derived peptides with broad spectrum antimicrobial activity. **Tack, Brian E.**; McCray, Paul; Welsh, Michael; Travis, Sue M.; Lehrer, Robert (University of Iowa Research Foundation, USA; The Regents of the University of California). PCT Int. Appl. WO 2001012668 A1 20010222, 137 pp. DESIGNATED STATES: W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM; RW: AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG. (English). CODEN: PIXXD2. APPLICATION: WO 2000-US22781 20000818. PRIORITY: US 1999-PV149886 19990818.

AB The invention relates to the use of antimicrobial peptides in the inhibition of microbial growth and proliferation. Antimicrobial truncated peptides are disclosed which are based on SMAP 29 and RCAP 18, but which contain a lesser no. of amino acid residues yet still retain bactericidal activity. In addn., synthetic peptides based upon the SMAP 29 protein are disclosed which have fewer amino acid residues and include substitutions yet retain substantial activity. The invention also relates to a method of inhibiting microbial growth by administering an effective amt. of a peptide in accordance with the invention, or by combining the peptides with other antimicrobial agents or antibiotics.

L73 ANSWER 2 OF 2 MEDLINE

DUPLICATE 1

2000231814 Document Number: 20231814. PubMed ID: 10768969. Bactericidal activity of mammalian cathelicidin-derived peptides. Travis S M; Anderson N N; Forsyth W R; Espiritu C; Conway B D; Greenberg E P; **McCray P B Jr; Lehrer R I; Welsh M J; Tack B F.** (Department of Internal Medicine, University of Iowa College of Medicine Iowa City, Iowa 52242, USA.) INFECTION AND IMMUNITY, (2000 May) 68 (5) 2748-55. Journal code: 0246127. ISSN: 0019-9567. Pub. country: United States. Language: English.

AB Endogenous antimicrobial peptides of the cathelicidin family contribute to

innate immunity. The emergence of widespread antibiotic resistance in many commonly encountered bacteria requires the search for new bactericidal agents with therapeutic potential. Solid-phase synthesis was employed to prepare linear antimicrobial peptides found in cathelicidins of five mammals: human (FALL39/LL37), rabbit (CAP18), mouse (mCRAMP), rat (rCRAMP), and sheep (SMAP29 and SMAP34). These peptides were tested at ionic strengths of 25 and 175 mM against *Pseudomonas aeruginosa*, *Escherichia coli*, *Staphylococcus aureus*, and methicillin-resistant *Staphylococcus aureus*. Each peptide manifested activity against *P. aeruginosa* irrespective of the NaCl concentration. CAP18 and SMAP29 were the most effective peptides of the group against all test organisms under both low- and high-salt conditions. Select peptides of 15 to 21 residues, modeled on CAP18 (37 residues), retained activity against the gram-negative bacteria and methicillin-sensitive *S. aureus*, although the bactericidal activity was reduced compared to that of the parent peptide. In accordance with the behavior of the parent molecule, the truncated peptides adopted an alpha-helical structure in the presence of trifluoroethanol or lipopolysaccharide. The relationship between the bactericidal activity and several physiochemical properties of the cathelicidins was examined. The activities of the full-length peptides correlated positively with a predicted gradient of hydrophobicity along the peptide backbone and with net positive charge; they correlated inversely with relative abundance of anionic residues. The salt-resistant, antimicrobial properties of CAP18 and SMAP29 suggest that these peptides or congeneric structures have potential for the treatment of bacterial infections in normal and immunocompromised persons and individuals with cystic fibrosis.

=> log y

COST IN U.S. DOLLARS	SINCE FILE	TOTAL
	ENTRY	SESSION
FULL ESTIMATED COST	27.26	213.52
DISCOUNT AMOUNTS (FOR QUALIFYING ACCOUNTS)	SINCE FILE	TOTAL
	ENTRY	SESSION
CA SUBSCRIBER PRICE	-4.96	-13.22

STN INTERNATIONAL LOGOFF AT 09:13:19 ON 12 JUL 2002